



TITLE:

プロテインホスファターゼ遺伝子族の多様化に関する分子進化学的解析(Dissertation_全文)

AUTHOR(S):

小柳, 香奈子

CITATION:

小柳, 香奈子. プロテインホスファターゼ遺伝子族の多様化に関する分子進化学的解析. 京都大学, 2000, 博士(理学)

ISSUE DATE:

2000-03-23

URL:

<https://doi.org/10.11501/3167158>

RIGHT:

学位申請論文

プロテインホスファターゼ遺伝子族の
多様化に関する分子進化学的解析

京都大学大学院理学研究科生物科学専攻

小柳香奈子

要旨

生物の進化を分子レベルから理解する目的で、生物の進化と遺伝子重複との関連性について、プロテインホスファターゼ遺伝子族を用いて分子系統学的解析を行った。生物の形態進化は、遺伝子重複により新しい遺伝子を次々と作り出すことで達成されたのだろうか。このことを明らかにする目的で、プロテインチロシンホスファターゼ(PTP)遺伝子族とプロテインセリン・スレオニンホスファターゼ(PP)遺伝子族の2つの遺伝子族について、遺伝子重複による多様化が起きた時期を明らかにした。

まず、現存する動物門中最古に分岐した原始的な多細胞動物である海綿動物から、RT-PCRによりPTP遺伝子の探査を試みた。その結果、海綿動物から10種類のPTP遺伝子の単離に成功した。さらに単離した遺伝子を含めた分子系統学的解析から、機能の異なるサブタイプの多様化が、動物界最古の種分岐である海綿動物とその他の動物の分岐前（10.7億年前～9.4億年前）に完了していたことがわかった。続いて、機能は同じだがしばしば発現する組織の異なるアイソフォームの多様化の時期を知る目的で、脊椎動物に最も近縁な無脊椎動物である頭索動物、脊椎動物中最古に分岐する無顎類、及び有顎類中最も原始的な軟骨魚類から、RT-PCRによりPTP遺伝子の探査を試みた。その結果、頭索動物から9種類、無顎類から11種類、軟骨魚類から13種類のPTP遺伝子の単離に成功した。これらの遺伝子を含めた分子系統学的解析の結果、アイソフォームの多様化は、有顎類と無顎類が分岐した5億年前ごろに集中して起きていることが明らかとなった。

次に、より幅広い生物種に存在しているPP遺伝子族と、真核生物の多様化がどう関連しているかを知る目的で、現存する真核生物中最古に分岐したミトコンドリアをもたない原始的な原生生物であるギアルディアのDNA配列データベースから、PP遺伝子の探査を行った。その結果、ギアルディアにも9種類のPP遺伝子が存在することがわかった。さらにこのギアルディアの遺伝子を含めた分子系統学的解析の結果、機能の異なるサブタイプの多様化が、真核生物最古の種分岐であるギアルディアとその他の真核生物の分岐前にほぼ完了していたことが明らかとなった。またアイソフォームの多様化した時期を知る目的で、PTP遺伝子族と同様の多様な動物すなわち海綿動物・頭索動物・無顎類・軟骨魚類からRT-PCRによりPP遺伝子の探査を試みた。その結果、海綿動物から6種類、頭索動物から4種類、無顎類から5種類、軟骨魚類から9種類のPP遺伝子の単離に成功した。これらの遺伝子を含めた分子系統学的解析から、アイソフォームの多様化はPTP遺伝子族同様、主に有顎類と無顎類が分岐した5億年前ごろに集中して起きていることが示された。

以上の結果から、脊椎動物へ至る系統でみると、遺伝子重複による遺伝子の多様化は、生物の進化にともない徐々に起きているのではなく、ある時期に限定して断続的に起きていることが明らかとなった。すなわち断続的な遺伝子重複がみられる時期は、A) 真核生物が核を獲得し細胞骨格や細胞内小器官などの細胞内構造が進化した時期、B) 動物が多細胞性を獲得した時期、C) 脊椎動物の神経系・消化器系などの独立した組織や器官が進化した時期、に対応する。一方、断続的な遺伝子重複が見られない時期に起きた生物の進化、例えば6～7億年前の多様な動物が爆発的に進化したカンブリア爆発と遺伝子重複とは関連が無く、そのような生物の進化は新しい遺伝子を創造することなく既存の遺伝子を利用して達成されたと考えられた。

目次

序論	3
1. 分子進化学の歴史	
2. 形態進化と分子進化	
3. 遺伝子重複と生物の進化	
4. これまでの研究	
5. 本論文の目的	
6. プロテインホスファターゼの分類と機能	
材料と方法	13
1. プロテインホスファターゼ遺伝子のクローニング	
2. 分子進化学的解析	
結果と考察	
1. プロテインチロシンホスファターゼ(PTP)遺伝子族	
1-1. PTP遺伝子族のサブタイプの多様化	26
結果	
考察	
1-2. PTP遺伝子族のアイソフォームの多様化	49
結果	
考察	

2. プロテインセリン・スレオニンホスファターゼ(PP)遺伝子族	
2-1. PP遺伝子族のサブタイプの多様化	73
結果	
考察	
2-2. PP遺伝子族のアイソフォームの多様化	89
結果	
考察	
総合議論	106
1. 結論のまとめ	
2. 遺伝子族の多様化に関する今後の課題	
3. 新しい遺伝子を創造することなく多様な生物が進化するしくみ	
4. 展望	
謝辞	114
文献	115
付録	123

序論

1. 分子進化学の歴史

分子進化学は、1965年のZuckerlandlとPaulingの分子時計に関する研究が発端となりスタートした(Zuckerlandl & Pauling, 1965)。分子進化学は、進化の証拠を化石ではなく分子情報、すなわち核酸やタンパク質の配列情報に求める学問であり、その最大の特徴は分子情報を取り扱うことで進化に関する客観的すなわち定量的・統計的な解析が可能なことである。そして、現在までに従来の進化学では明らかにされなかった多くの事実を解明してきた。

一般に進化学における研究は大きく二つの方向に分類される(Futuyma, 1998)。一方は進化の機構を分析する研究（機構の研究）であり、もう一方は個々の生物集団と生命の歴史を推定する研究（歴史の研究）である。分子進化学も例外ではない。分子進化学における機構研究の最も重要なものに木村の中立説がある。中立説は「分子レベルでは、淘汰に有利でも不利でもない中立な変異が集団に広まった結果、進化が起こる」と主張するもので、分子レベルでは形態レベルと進化の機構が異なる（後述）、という事実を明らかにした(Kimura, 1983)。また分子進化学における歴史の研究は、化石記録からは推定が困難であった生物の歴史を明らかにした。例えば、1977年Woeseらは初めて分子を用いて、バクテリアを含む全生物の系統分類を行った(Woese & Fox, 1977)。近年では様々な分類群に属する生物の系統関係が、分子系統樹を用いて推定されている。また最近では、生物の分岐年代を推定するという研究の重要性も増している(e.g. Doolittle et al., 1996; Wray et al., 1996; Nikoh et al., 1997)。このように、近年の分子生物学の発展及び生命科学技術の向上に伴う分子情報の蓄積に後押しされ、分子進化学は大きな成果をあげてきた。

しかし依然として、分子進化学が当初から掲げていた問題は未解決である(Kimura, 1991; Miyata et al., 1994)。すなわち「生物の進化の分子的理解」という問題が、未解決のまま残されているのである。この問題は、進化の機構と歴史という二つの研究の流れを統合するものであり、分子進化学に残された最大のテーマだといえる。本論文はこの「生物の進化の分子的理解」をテーマに行ったものである。

2. 形態進化と分子進化

「生物の進化の分子的理解」は難題である。なぜなら現在、形態進化と分子進化は異なる機構で起きる、と理解されているからである。形態の進化は、自然淘汰により環境に対して有利な形質をもつ個体が子孫を多く残し、その形質が集団に広まった結果、適応的に起きる、と考えられている(Futuyma, 1998)。実際にダーウィン・フィンチの嘴の進化などは、適応的な進化の具体例とされる(Grant, 1986)。一方分子の進化は、有利でも不利でもない中立な突然変異をもつ遺伝子が、偶然に集団に広まった結果、中立的に起きる、と考えられている(Kimura, 1983)。形態の進化で重要だと考えられている有利な突然変異は、分子レベルでは無視できるほど僅かだとされる。分子進化速度の一定性や偽遺伝子の進化速度の上昇が、このことを証明している(Kimura, 1983; Kimura, 1991)。このように「形態の進化」と「分子の進化」を支配する機構の間には隔たりがある。しかし、個体のもつ形質の基盤は分子が決定していると考えられるから、「形態の進化」と「分子の進化」との間には何らかの関連があるはずである。それでは「形態の進化」と「分子の進化」との間に存在するギャップを埋めるにはどうしたらよいか。すなわち、生物の形態進化を分子的に理解するためにはどうしたらよいか。

3. 遺伝子重複と生物の進化

「生物の進化の分子的理解」という問題は、「生物の形態進化にともなって遺伝子に何が起きたか」という問題に置き換えることができる。「遺伝子に何が起きたか」といった場合には、遺伝子の突然変異・遺伝子の重複・ゲノムの再編成などが考えられるが、中立説の基盤となった突然変異以外の要因に注目する必要があるだろう。そこで本論文では遺伝子重複に注目した。遺伝子重複は、既存の遺伝子の機能を保持したまま、新しい機能をもった遺伝子を作り出す機構である(Ohno, 1970)。すなわち遺伝子重複が起きてコピーが出来ると、一方が従来の機能を果たしながら、もう一方は突然変異を蓄積し新しい機能をもった遺伝子へと進化することが出来る。こうしてできた新しい遺伝子が、生物の新しい体制や機能の獲得に関わった可能性は十分に考えられる(Ohno, 1970; Ohta, 1991)。いいかえれば、生物の形態進化にともなって、遺伝子重複により新しい遺伝子が次々と作り出されている可能性がある。したがって、生物の進化にともなって遺伝子重複により新しい遺伝子が作り出されているかどうか、作り出されているとしたらそれはどのような特徴をもった遺伝子なのか、を解明することにより生物の進化を分子的に理解する手がかりが得られると考えられる。

4. これまでの研究

岩部ら(Iwabe et al., 1996)は25の遺伝子族について、生物進化の過程において、いつ遺伝子重複が起きて新しい遺伝子が作り出されたのか、そしてそれはどのような特徴をもつ遺伝子か、分子進化学的解析を行った。遺伝子族とは、共通の祖先配列から遺伝子重複が繰り返された結果作り出された遺伝子グループのことである。例えば、よく知られた遺伝子族にプロテインキナーゼ遺伝子族がある。この遺伝子族に属する遺伝子には、cAMP依存性キナーゼ(PKA)やカルモジュリン依存性キナーゼなど、調節を受ける分子の異なる遺伝子や、PKAの中でもPKA α やPKA β など発現する組織の異なる遺伝子など(Taylor et al., 1990 for review)、多様な遺伝子が存在する。岩部らは25の遺伝子族についてどの遺伝子がいつ作り出されたのかを調べた。同じ遺伝子族に属する遺伝子は、遺伝子配列内に互いに相同性のある領域が存在する。したがって、各遺伝子族について分子系統樹を推定することが可能である。岩部らは、様々な生物由来の遺伝子を含めて分子系統樹を推定することで、遺伝子重複により遺伝子が分かれた時期を推定した。その結果、脊椎動物に至る進化の過程において、遺伝子重複が起きて多様な遺伝子が作り出された時期が、三つの時期に分類されることを明らかにした。一つめは原核生物（真正細菌及び古細菌）と真核生物の分岐後、動物・植物・菌類の分岐前の時期である（図1のA期）。この時期には、イオンポンプや細胞骨格タンパク質の遺伝子族で遺伝子重複が起き、細胞内局在の異なる遺伝子が作り出された。二つめは動物・植物・菌類の分岐後、旧口動物と新口動物の分岐前の時期である（図1のB期）。この時期には、シグナル伝達に関わる遺伝子族で遺伝子重複が起き、機能の異なる遺伝子が作り出された。三つめは旧口動物と新口動物の分岐後、硬骨魚類と四足動物の分岐前の時期である（図1のC期）。この時期には、様々な遺伝子族で遺伝子重複が起き、発現する組織の異なる遺伝子が作り出された。岩部らは、遺伝子重複が起きて多様な遺伝子が作り出された三期すなわちA期、B期、C期は、生物の形態上大きな進化がみられた時期に対応するのではないか、と考察している。すなわち、遺伝子重複による遺伝子族の多様化と、真核生物の初期進化（A期）、動物における多細胞性の進化（B期）、脊椎動物の器官や組織の進化（C期）、の間に関連性のあることが示唆された。

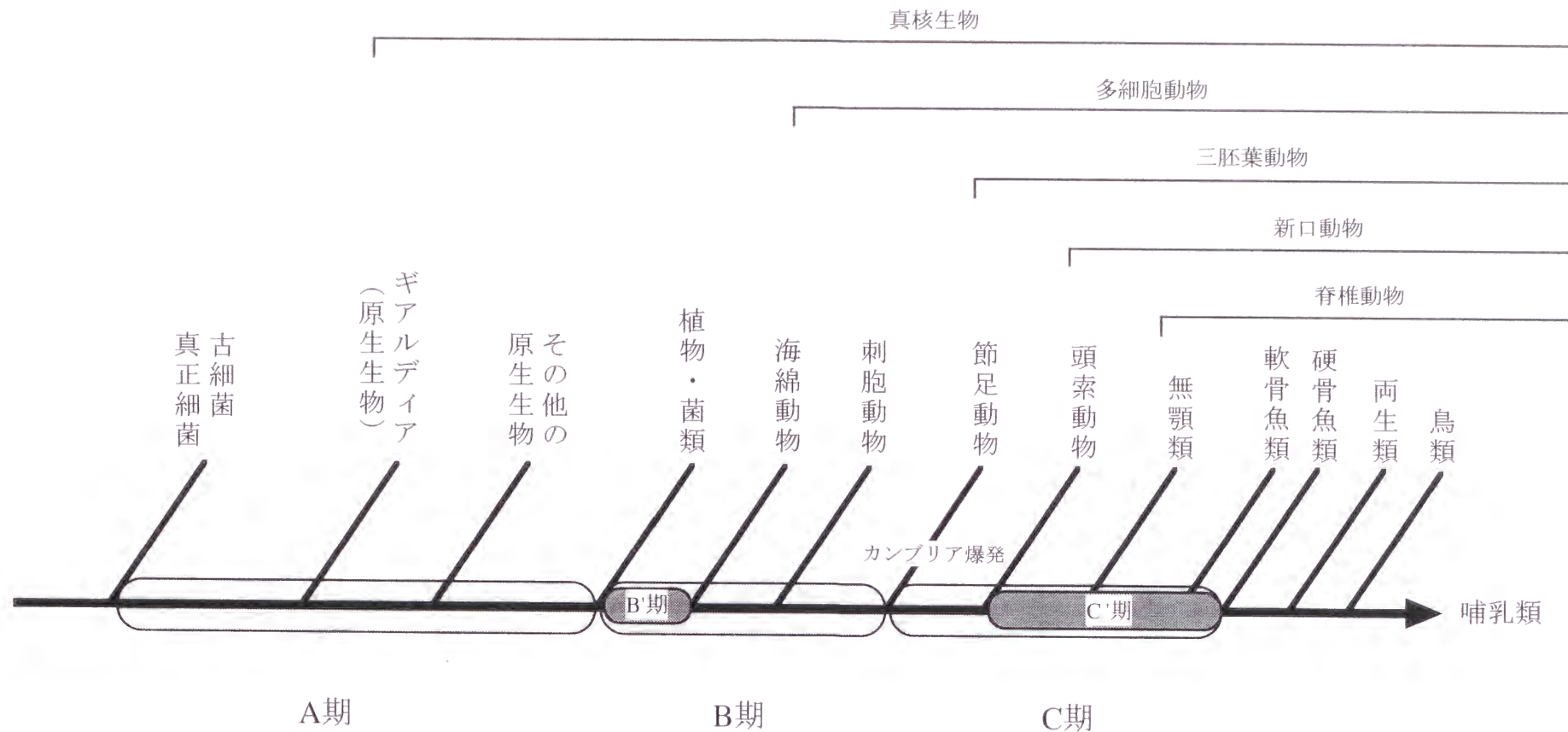


図1 これまでの研究

これまでの研究成果をもとに、遺伝子重複が起きた時期と生物の進化の対応を模式的に示した。遺伝子重複が起きた時期は網掛けをして示した。生物の系統関係はWillmer, 1990; Romer & Parsons, 1986に基づいた。真正細菌・古細菌・真核生物の関係については、古細菌と真核生物が近縁であることが示されているが(Iwabe et al., 1989)、真正細菌・古細菌・真核生物間で頻繁に遺伝子の水平移動が起きたことが知られているため(Doolittle & Brown, 1994)、図では真正細菌・古細菌・真核生物が一斉に分岐したように記した。また、ギアルディアは真核生物の中で最も古くに分岐した原生生物だと考えられている(Sogin et al., 1989; Hashimoto et al., 1994)。ギアルディアの分岐後には、多様な原生生物が分岐したと考えられるが(Cavalier-Smith, 1993)、図では「その他の原生生物」としてまとめて示した。枝の長さは概ね時間を反映しているが正確ではない。正確な分岐年代については材料と方法の章参照。

さらに菅ら(Suga et al., 1999a)は、Gタンパク質 α サブユニット($G\alpha$)遺伝子族及びプロテインチロシンキナーゼ(PTK)遺伝子族を用いて、B期に起きた遺伝子重複について詳細な解析を行った。菅らは遺伝子重複の起きた時期を限定する目的で、最も原始的な多細胞動物である海綿動物及び二胚葉動物である刺胞動物から $G\alpha$ 遺伝子とPTK遺伝子をクローニングし、分子進化学的解析を行った。その結果、 $G\alpha$ 遺伝子族・PTK遺伝子族とも、遺伝子重複が起きて機能の異なる遺伝子の作り出された時期が、菌類と動物の分岐後、海綿動物とその他の動物の分岐前という、多細胞動物進化の初期のごく限定された時期であることを示した(図1のB'期)。同様の結果が、ホスホジエステラーゼ(PDE)遺伝子族(Koyanagi et al., 1998a)、ホスホリパーゼC(PLC)遺伝子族・プロテインキナーゼC(PKC)遺伝子族(Koyanagi et al., 1998b)、*Pax*遺伝子族(Hoshiyama et al., 1998)でも得られた。

また菅ら(Suga et al., 1999b)は、PTK遺伝子族を用いてC期に起きた遺伝子重複についても詳細な解析を行った。菅らは、脊椎動物に最も近い無脊椎動物である頭索動物(Willmer, 1990; Wada and Satoh, 1994)、脊椎動物の中で最も原始的とされる無顎類(Romer & Parsons, 1986)からPTK遺伝子をクローニングし、分子進化学的解析を行った。その結果、遺伝子重複により発現する組織の異なる遺伝子の作り出された時期が、無顎類の分岐の周辺という、脊椎動物進化の初期であることを示した(図1のC'期)。

以上の結果から、遺伝子重複はB'期及びC'期という限定した時期に起きていることが明らかとなった。それぞれの時期は、動物の初期進化において多細胞性が獲得された時期、脊椎動物の初期進化において器官や組織の進化が起きた時期に対応する。したがって、岩部らの示唆するように、これらの遺伝子重複と生物の進化との間には関連性のある可能性がある。しかし一方で、B'期やC'期以外には遺伝子重複が起きていないことが明らかとなった。したがって、B'期やC'期以外に起きた生物の進化、例えば多様な動物門が分岐したカンブリア爆発と遺伝子重複とは、関連の無いことが示唆された(Suga et al., 1999a; 1999b; Koyanagi et al., 1998a; 1998b; Hoshiyama et al., 1998)。

5. 本論文の目的

これまでの研究から得られた結論は、生物の形態進化を分子レベルから理解するために重要なものであるが、数遺伝子族により得られた結果であり、さらに多くのデータを蓄積し検証する必要がある。また遺伝子重複が起きた時期をさらに限定し、生物の形態進化と遺伝子重複の関係をより密接に関連づけて考察する必要がある。そこで本論文では、プロテインホスファターゼ遺伝子族について、遺伝子重複により新しい遺伝子が作り出された時期はいつか、その結果どのような特徴をもつ遺伝子が作り出されたのか、分子進化学的解析を行った。

タンパク質のリン酸化・脱リン酸化は、細胞のシグナル伝達を通じて転写・翻訳・細胞周期・代謝などの制御を行い、ひいては細胞の運命すなわち増殖・分化などの決定に関わる(Alberts et al., 1994)。このタンパク質のリン酸化の状態は、リン酸化を行うプロテインキナーゼと、脱リン酸化を行うプロテインホスファターゼが、表裏一体となり決定している。これらプロテインキナーゼやプロテインホスファターゼの多様性は、細胞内のシグナル伝達システムの複雑性に寄与するから、さらに高次のレベルにおける生命現象の複雑性にも関与すると考えられる。すなわち、タンパク質のリン酸化・脱リン酸化を制御する遺伝子の多様化は、生物の進化に深く関与した可能性がある。この点から、プロテインホスファターゼは本論文のテーマに適した分子だといえる。

真核生物のタンパク質においては、主にチロシン残基・セリン残基・スレオニン残基がリン酸化を受ける(Alberts et al., 1994)。本論文では、チロシン残基を脱リン酸化する酵素であるチロシンホスファターゼ(PTP)遺伝子族と、セリン残基・スレオニン残基を脱リン酸化する酵素であるセリン・スレオニンホスファターゼ(PP)遺伝子族の二遺伝子族について解析を行った。PTP遺伝子は、動物・植物・菌類・細胞性粘菌に存在しており、特に動物においては多様なPTP遺伝子が存在することが知られている(Streuli, 1996; Charbonneau & Tonks, 1992 for review)。これに対して、PP遺伝子は古細菌と真核生物に広く存在しており、真核生物において多様なPP遺伝子が存在することが知られている(Shenolikar, 1994; Mumby & Walter, 1993 for review)。

これまでの研究から、遺伝子族の遺伝子重複による多様化の起きた時期は、三つの時期(A期, B'期, C'期)に分類されることが示された。それではPTP遺伝子族及びPP遺伝子族の多様な遺伝子も、この三期に作り出されたのだろうか。本論文ではこのことを明らかにする目的で、時期推定に重要だと考えられる複数の生物種からプロテインホスファターゼ遺伝子のクローニングを行い、分子進化学的解析を行った。特に本論文においては、A、

B', C'の各期について、以下の点に注目しより詳細な解析を行った。

B'期には、 $G\alpha$ ・PTK・PDE・PLC・PKC・*Pax*遺伝子族で遺伝子重複による多様化が見られ、このことと動物の多細胞化との関連が示唆されている。遺伝子重複と多細胞化との関連を考える上で特に、受容体型のPTP遺伝子はよい材料となる。なぜなら、受容体型のPTP遺伝子の中には細胞接着に密接に関わる遺伝子があり(Brady-Kalnay & Tonks, 1995; Schaapveld et al., 1997 for review)、多細胞性の獲得に重要だったと考えられるからである。本論文では、原始的な多細胞動物である海綿動物から、受容体型PTP遺伝子の単離を試みた。

C'期に起きた遺伝子重複については、PTK遺伝子族一つでしか詳細な解析がされていない。PTK遺伝子族の結果から、遺伝子重複が起きた時期は無顎類と有顎類の分岐周辺であることが明らかとなったが、本論文ではその時期が無顎類と有顎類の分岐以前であったのか以降であったのか明らかにすることを試みる。また、さらにその時期を限定する目的で、軟骨魚類に注目した。これまでの結果から遺伝子重複が起きた時期は、少なくとも硬骨魚類と四足動物の分岐前であることが示されているが、硬骨魚類よりも古くに分岐するとされる軟骨魚類(Romer & Parsons, 1986; Colbert & Morales, 1991)の分岐以前であったのか以降であったのかについてはわかっていない。本論文ではこの問題を明らかにした。

A期に起きた遺伝子重複については、これまで詳細な解析がされていない。本論文では、真核生物の中で最も古い時期に分岐する原生生物と考えられているギアルディア(Sogin et al., 1989; Hashimoto et al., 1994)のDNA配列データベース(材料と方法の章参照)を用いて、遺伝子重複の起きた時期の限定を試みた。尚、PTP遺伝子族についてはこの時期に多様化がみられないことから、A期についてはPP遺伝子族についてのみ解析を行った。

6. プロテインホスファターゼの分類及び機能

プロテインホスファターゼの研究は、プロテインキナーゼに比べて遅れていた。しかし、最近になりプロテインホスファターゼの研究は急速に進み、その役割が明らかとなっている(Hunter, 1995; Zolnierowicz & Hemmings, 1996; 武田、1998 for review)。当初、タンパク質のリン酸化状態は特異性の高い膨大な種類のプロテインキナーゼによって決定され、プロテインホスファターゼは基質特異性がきわめて広く絶えず活性をもつ状態で存在している、とする説があった。しかしその後、プロテインホスファターゼにも多くの種類が存在し、積極的にタンパク質機能の調節に関わり、細胞機能に不可欠な役割を果たしていることが明らかとなった。例えば、ショウジョウバエの受容体型のPTPであるDLARは、神経系の発生において、神経細胞の標的細胞の認識に必要であることが、変異体の解析から明らかとなった(Chien, 1996; Stoker & Dutta, 1998 for review)。また、マウスの細胞質型のPTPであるSHP2は、胚発生の初期段階で重要な役割を果たすことが、ノックアウトマウスの解析から明らかとなっている(Saxton et al., 1997)。さらに、PTPはPTKと比べて活性が非常に高いため、その機能が非常に厳しく制御される必要性のあることも指摘されている(Hunter, 1995)。このように研究の進展にともない、プロテインホスファターゼの重要性が明らかとなってきた。

真核生物のタンパク質においては、主にチロシン残基・セリン残基・スレオニン残基がリン酸化を受ける(Alberts et al., 1994)。プロテインホスファターゼは、この脱リン酸化するアミノ酸の特異性により3つに分類される(Hunter, 1995; Zolnierowicz & Hemmings, 1996)。すなわち、プロテインチロシンホスファターゼ(PTP)、プロテインセリン・スレオニンホスファターゼ(PP)、プロテインセリン・スレオニン・チロシンホスファターゼ(二重特異性ホスファターゼ、DSP)、である(表1)。さらにPPは、PP1やPP2Aなどの主要な分子の属するPP(あるいはPPP)遺伝子族と、 Mg^{2+} イオンに依存性がありPP遺伝子族とは相同性のないPPM遺伝子族とに分類される。このうち本論文では、大きな遺伝子族を成すPTP遺伝子族とPP遺伝子族の二遺伝子族をとりあげる。PTP遺伝子とPP遺伝子の間に相同性は認められず、両者の脱リン酸化反応機構も全く異なっている(Denu et al., 1996)。したがって二遺伝子族は異なる起源をもち、異なる進化の過程を経て、別々の遺伝子族を形成したと考えられる。

動物のPTP遺伝子は、細胞の増殖・分化・接着の制御を通じて、発生や造血細胞の分化などの生命現象に関わっている(Charbonneau & Tonks, 1992; Hunter, 1995; Streuli, 1996 for review)。植物においては、発生や環境応答などの生命現象に関与することが示唆されてい

る(Xu et al., 1998)。また菌類においては、細胞分裂の制御などの機能を果たしている(e.g. Otilie et al., 1992)。一方PP遺伝子は、幅広い生物すなわち古細菌と真核生物に存在している。PP遺伝子は、主に代謝、細胞周期、転写の調節など細胞の基本的な機能の調節を行うことが知られている(Shenolikar, 1994; Mumby & Walter, 1993 for review)。

表1 プロテインホスファターゼ遺伝子族の分類

脱リン酸化を受けるアミノ酸	遺伝子族	ヒトの主な遺伝子
チロシン残基	PTP*	LCA, LAR, SH-PTP1
セリン・スレオニン残基	PP* (PPP)	PP1 α , PP2A α , PP2B α
	PPM	PP2C α , PP2C β
チロシン残基及びセリン・スレオニン残基	VH	CL100, VHR
	Cdc25	Cdc25A, Cdc25B, Cdc25C

Hunter, 1995; Zolnierowicz & Hemmings, 1996に基づき、 プロテインホスファターゼ遺伝子族の分類をまとめた。本論文で解析を行う遺伝子族に*を記した。

材料と方法

1. プロテインホスファターゼ遺伝子のクローニング

生物試料の入手

カワカイメン、ナメクジウオ、ヌタウナギ及びポタモトリゴンモトロ（淡水エイ）から、チロシンホスファターゼ(PTP)遺伝子及びセリン・スレオニンホスファターゼ(PP)遺伝子の探査を行った。カワカイメンは、茨城県霞ヶ浦において採集されたものを、本研究室で無機培養した（後述）。ナメクジウオは、愛知県渥美半島沖の海底から採集されたものを、東京大学海洋研究所の窪川かおる博士より提供を受けた。ヌタウナギは、神奈川県三崎海岸沖で採集されたものを、東京大学理学部附属臨海実験所より提供を受けた。淡水エイは熱帯魚販売店（京都熱帯魚）より購入した。四種の生物の和名、生物種名、学名及び分類群を表2に示した。

表2 材料に用いた生物種の和名、生物種名、学名及び分類群

和名	生物種名 (common name)	学名	分類群
カワカイメン	sponge	<i>Ephydatia fluviatilis</i>	海綿動物門、普通海綿綱
ナメクジウオ	amphioxus	<i>Branchiostoma belcheri</i>	脊索動物門、頭索動物亜門、 ナメクジウオ綱
ヌタウナギ	hagfish	<i>Eptatretus burgeri</i>	脊索動物門、脊椎動物亜門、 無顎動物下門、メクラウナギ綱
ポタモトリゴンモトロ	ray	<i>Potamotrygon motoro</i>	脊索動物門、脊椎動物亜門、 軟骨魚綱

分類群は岩波生物学辞典第4版(1996)に基づいた。

RNAの抽出とcDNAの作成

カワカイメンは生育条件の悪化に対応するために、芽球(gemmule)と呼ばれる休眠構造を作る。この芽球は、細胞が多数集合して球状塊をつくり、その外側を硬い殻が覆ったものである(Harrison & Cowden, 1982)。芽球を用いることでカワカイメンの無機培養が可能であり、他生物の混入の危険性を抑えることができる。そこで、試料として芽球を発芽させたものを用いた(Rasmont, 1961; Seimiya et al., 1994)。

ナメクジウオは生体を超純水で洗浄した6個体を用いた。

ヌタウナギ及びエイに関しては、70%エタノールで麻酔・洗浄後、解剖を行った。他生物の混入の危険性のある消化管は除き、ヌタウナギにおいては脳・肝臓・眼（痕跡器官）を、エイにおいては脳・肝臓・眼・卵巣・脾臓を摘出した。

以上の各試料をホモジェナイズし、TRIZOL(GIBCO BRL)又はセパゾールRNA I(nacalai tesque)を用いてRNAを抽出した。抽出したtotal RNAは、5'末端にアダプター配列のついたoligo-dTプライマー(5'-GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTTTTTT-3')を用い、mRNAを選択的に逆転写しcDNAを作成した(3' RACE System for Rapid Amplification of cDNA Ends, GIBCO BRL)。

degenerate PCR法によるホスファターゼ遺伝子の探査

cDNAを鋳型としてdegenerate primerを用いたPCRにより、プロテインホスファターゼ遺伝子断片の増幅を行った。degenerate primerの設計は以下の手順で行った。まず、既知のPTP遺伝子及びPP遺伝子を問い合わせ配列として、既存のデータベース（後述）に対して相同性検索（後述）を行い、データベースからPTP遺伝子及びPP遺伝子のデータを抽出した。抽出されたPTP遺伝子配列群及びPP遺伝子配列群をそれぞれアライメント（後述）し、配列全体でアミノ酸が連続して保存している領域を複数箇所選択した。そして、選択された領域で保存しているアミノ酸をコードし得る全ての塩基配列を含むようにdegenerate primerを設計した（表3）。

表3のdegenerate primerを用いPCR反応を行った。サーマルサイクラーにはGeneAmp PCR System2400, 9600, 9700(PE Applied Biosystems)を、反応液にはExpand High Fidelity PCR System(Roche)あるいはAmpli Taq Gold(PE Applied Biosystems)を用いた。目的の長さの断片が得られなかった場合には、最初のPCR反応液を鋳型として、最初に用いたdegenerate primerの内側に位置する別のdegenerate primerを用いてnested PCRを行った。PCRの反応温度、時間及びサイクル数は、94度で2分(Expand High Fidelity PCR System)または9分(Ampli Taq Gold)後、94度0.5-1分、45-46度1-2分、72度2-5分を5-10回に続いて、94度0.5分、60度1

分、72度2分を20-30回、60度5分、72度10分とした（幅のある値は、増幅するDNA断片の長さ、degenerate primerのmelting temperature等に応じる）。各遺伝子について用いたdegenerate primerの詳細な組み合わせは付録1-1に示した。

表3 PTP遺伝子及びPP遺伝子の単離に用いたdegenerate primer

primer名		塩基配列 (5'側→3'側)	アミノ酸配列 (N側→C側)
PTP遺伝子			
S	ptp-1	CAGGATCC-TYTGGMGNATGRTNTGG	FWRM(I/V)M
	ptp-2	CAGGATCC-INGAYTTYTGGMGNATG	XDFWRM
A	ptp-a	GTGAATTC-RYICCIGCNSWRCARTG	HCSAG(V/T/A)
	ptp-b	GAGAATTC-GTICKNCCNACNCCNGC	AGVGRT
PP遺伝子			
S	pp-1	CACTGTCG-GGIGAYNYNCA YGGNCA	GD(I/V/L/T)HGQ
	pp-2	GCTGTTCA-WIGGNGAYTWYGTNGAY	(L/M/N)GD(Y/F)VD
	pp-3	CTCATGGA-GAYTWYGTNGAYMGNGG	D(Y/F)VDRG
A	pp-a	CTGATAGG-TARTTIGGNGCNSWRAA	YNPASF

塩基配列はIUBコードで示した。Iはイノシンを表す。各プライマーの5'末端側には、PCR反応の後半においてアニーリング温度を上げ目的の遺伝子を得やすくする目的で8塩基を付加した。付加する8塩基については、プライマー全体のGC含量が50%前後になること、プライマーの5'末端側の上流で比較的保存しているアミノ酸配列をコードする塩基配列を含むこと、を考慮した。PTP-a, PTP-bはPTPの活性中心のアミノ酸配列を含んでいる。Sはセンス側のプライマー、Aはアンチセンス側のプライマーである。

PTP遺伝子族に関しては、表3のdegenerate primerを用いても特定のサブタイプ（サブタイプの定義は結果と考察1-1参照）が得られない場合があった。その場合には、ある特定のサブタイプが増幅されるように、サブタイプに特徴的な配列を含むdegenerate primerを作成した。すなわち、あるサブタイプのみでアミノ酸が連続して保存しており、その他のサブタイプではそのアミノ酸が保存していないような領域を選択し、選択された箇所degenerate primerを作成した（表4）。各遺伝子について用いたdegenerate primerの詳細な組み合わせは付録1-2に示した。

表4 PTP遺伝子の特定のサブタイプの単離に用いたdegenerate primer

サブタイプ	primer名		塩基配列 (5'側→3'側)	アミノ酸配列 (N側→C側)
PTPR2A	S	R2A-1	ACGGNCARCANTTYACNTGG	GQ(Q/H)FTW
PTPR4	A	R4-a	GTGAATTC-ADYTCNGTRTCNCCRTA	YGDTE(L/I)
PTPR5	S	R5-1	TGCCGATG-AAISANTTYSYNAARCA	(K/N)(H/Q/E)F(P/V/A)KH
		R5-2	CAGGATCG-MANCAYCCNGANAAYAA	(N/Q)HP(D/E)NK
	A	R5-a	GTGAATTC-CKYTGISWICKDATRTG	HIR(S/T)QR
PTPN3	S	N3-1	GTGCATTC-AAYMGNTAYMGNGAYGT	NRYRDV
	A	N3-a	GTGAATTC-GTYTGDATNARNCCCAT	MGLIQT
PTPN6	S	N6-1	CAGTAGGC-TTYTGGGARGARTTYGA	FWEEFE
	A	N6-a	TGTCGCGG-ATIAYRTCDATNARDAT	ILID(I/V)I
		N6-b	CGATGTAC-TGYTGNACNGCCATRTA	YMAVQ(Q/H)

塩基配列はIUBコードで示した。Iはイノシンを表す。各プライマーの5'末端側には、PCR反応の後半においてアニーリング温度を上げ目的の遺伝子を得やすくする目的で8塩基を付加した（R2A-1を除く）。付加する8塩基については、プライマー全体のGC含量が50%前後になること、プライマーの5'末端側の5'側で比較的保存しているアミノ酸配列をコードする塩基配列を含むこと、を考慮した。サブタイプ名は結果及び考察の章1-1を参照。Sはセンス側のプライマー、Aはアンチセンス側のプライマーである。

PCRで増幅したDNA断片のクローニング及び遺伝子配列の決定

PCRで増幅したDNA断片は、Gene Clean II(BIO 101 Inc.)あるいはQIAquick Gel Extraction Kit(QIAGEN)を用いて精製した。精製されたDNA断片は、Ready-To-Go T4 DNA Ligase (Pharmacia Biotech)あるいはDNA Ligation Kit Ver.2 (TaKaRa)を用いて、プラスミド pT7Blue(Novagen)に組み込んだ。DNA断片が組み込まれたプラスミドによって、コンピテント細胞(DH5 α , TOYOBO)を形質転換させた。プラスミドを保持したコロニーに対して、プラスミドのマルチクローニングサイトの両端に相補的な内向きのプライマー対を用いてPCR反応を行い、プラスミド内に目的のDNA断片が入っているコロニーを選択した。インサートが目的の長さのDNA断片であった場合には、塩基配列の決定を行った。シーケンサはABI PRISM 377(PE Applied Biosystems)あるいは島津DNAシーケンサDSQ-1000L(SHIMADZU)を用いた。シーケンサとしてABI PRISM 377を使用した場合には、ABI PRISM dRhodamine Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems)あるいはABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems)を用い、ダイターミネーター法により塩基配列を決定した。また、シーケンサ

として島津DNAシーケンサDSQ-1000Lを使用した場合には、Thermo Sequenase fluorescent labelled primer cycle sequencing kit with 7-deaza-dGTP (Amersham)を用い、ダイプライマー法により塩基配列を決定した。

3'RACE法による3'末端側の遺伝子配列の決定

DNA断片の遺伝子配列を決定し相同性検索（後述）を行った結果、そのDNA断片がプロテインホスファターゼ遺伝子であった場合には、3'RACE法(Frohman et al., 1988)によりその遺伝子の3'末端側の塩基配列を、上記と同様の手段により決定した(3' RACE System for Rapid Amplification of cDNA Ends, GIBCO BRL)。

その他

cDNA作成時の逆転写反応のエラー、PCR反応時の複製エラー及び種内多型を考慮して、最終的な遺伝子配列は独立の複数のクローンのコンセンサス配列とした。選択的スプライシングバリエーションと思われるような複数種類のクローンが存在した場合には、他の生物のホモログと思われる遺伝子と最も広い領域にわたって相同性の認められるクローンを選んだ。

他の生物種の混入には細心の注意を払い、同時期に2種類以上の生物種は扱わなかった。

特に記載のない実験手順についてはSambrook et al.(1989)によった。

2. 分子進化学的解析

遺伝子配列データベース

遺伝子配列データベースとして、DDBJリリース39、PIRリリース61を利用した。尚、線虫の遺伝子は分子進化速度が速く、特にゲノムプロジェクトで配列の決定されたものの中にはエキソンが正しく推定されていないと思われる配列が存在した。このような遺伝子を分子系統樹推定に含めると、遺伝子重複を実際よりも多く見積もる危険性がある。そのため、結論が変わらないことを確認した上で、本論文では線虫の遺伝子は除いて分子進化学的解析を行った。

この他にギアルディアの遺伝子配列データベースとして、現在進行中の*Giardia lamblia*のゲノムプロジェクト(Smith et al., 1998)で公開しているデータベース(http://evol3.mbl.edu/GiardiaHTML/giardia_data.html)を利用した。これは、ショットガン法によるゲノムDNA断片配列のデータベースである。現在28328本の断片配列が公開され、断片配列の平均長は884ヌクレオチドである。本論文では、公開されているDNA断片配列を、ソフトウェアPHRAP(http://www.genome.washington.edu/UWGC/analysis_tools/phrap.htm)により連結・再構成した。したがって最終的な遺伝子配列は、複数のDNA断片配列から成る。このギアルディアのDNA配列データベースは予備的に公開されたもので、配列内にエラーが含まれる可能性が指摘されている(http://evol3.mbl.edu/GiardiaHTML/giardia_data.html)。そこで本論文においては、配列の信頼度を高めるために、以下の修正を加えたものを最終的な遺伝子配列とした。すなわち、複数のDNA断片配列間で塩基の食い違う座位は、多数決により塩基を決定し、最終的な遺伝子配列とした。多数決で決定出来ない場合には、分子系統樹推定などの解析において、その座位は除外した。欠失・挿入が見られる場合には、他の生物のホモログと思われる遺伝子と最も広い領域にわたって相同性の認められる方を選んだ。また、現在のところ、ギアルディアにはイントロンは存在しないと考えられている(Adam, 1991 for review)。したがって最終的な遺伝子配列を翻訳したものが、その遺伝子のコードするタンパク質の推定アミノ酸配列だと考えられる。本論文ではこのアミノ酸配列を解析に用いた。

相同性検索（ホモロジーサーチ）

相同性検索を行うソフトウェアとして、FASTA version3.0(Pearson & Lipman, 1988)及びBLAST version 2.0(Altschul et al., 1997)を用いた。ホモロジーのある配列全てをデータベースから抽出する際にはFASTAを利用した。またBLASTには、FASTAに比べてアライメントの際にギャップを考慮しないために、ホモロジー検出能力は劣るが高速なプログラムがある。そこで、塩基配列を決定した配列に対して、ホモロジーのある配列がデータベース中に存在するか否かを確認する際にはBLASTを利用した。

ドットマトリクス

二配列間のホモロジーの有無の確認には、ドットマトリクス法(Toh et al., 1983)を利用した。また、遺伝子内における繰り返し構造の検出にも用いた。

遺伝子配列群のアライメント

アライメントとは配列間のホモロジーが最大になるように、配列中に適宜ギャップを挿入して、配列の整列を行うことである。二配列のアライメントはダイナミックプログラミング法(Needleman & Wunsch, 1970)により行った。多数本の配列のアライメントはKatohの方法により行った(Katoh et al., manuscript in preparation)。この方法はFeng & Doolittle(1987)の方法に改良を加え、高速なアルゴリズムを適用することで計算量の問題を改善したものである。さらに目視による修正も行った。分子系統樹の推定には、一義的にアライメントされる領域のみを用い、一配列でもアミノ酸の欠失のみられる座位は除いた。

分子系統樹の推定

分子系統樹の推定方法は、推定に用いる配列の本数に応じて使い分けた。

A. 網羅的探索による最尤法

最尤法(Kishino et al., 1990; Adachi & Hasegawa, 1996)は、考えられる系統樹の樹形を網羅的に探索して、尤度（データが実現する確率）が最大となる系統樹を求める方法である。この方法は、仮定したアミノ酸置換のモデルが現実と違った場合にも頑強であることが知られている(Hasegawa & Fujiwara, 1993)。また、異なる樹形間の統計的な比較が可能である。本論文では、アミノ酸置換の確率法則としてJTTモデル(Jones et al., 1992)を用い、ソフトウェアとしてはMOLPHY2.3β3(Adachi & Hasegawa, 1996)を用いて、最尤系統樹推定を行った。

系統樹の確からしさは、RELL法(Kishino et al., 1990)に基づき、系統樹上の各分岐点において分岐点以下の配列が単系統となるブートストラップ確率(Felsenstein, 1985)を求めて検証した。

B. 発見的探索による最尤法

最尤法は系統関係を推定したい配列の数（系統樹における操作上の分類単位: OTU, Operational Taxonomic Unit）の増加に伴いその計算量が膨大になり、OTUが10以上の場合には、最尤系統樹の網羅的探索が事実上不可能となる。そこで、OTUが10以上の場合には発見的探索により最尤系統樹を推定した。本論文における発見的探索は、以下の手順で行った。まず近隣結合法（後述）により系統樹を推定する。これを初期系統樹とし、Tree Bisection and Reconnection(TBR, Swofford et al., 1996)による変形をほどこす。TBRにより得られる全ての変形された系統樹に対して尤度を計算し、その中で最も尤度の高い系統樹を求める。この最尤系統樹と初期系統樹の尤度を比較し、尤度が改善されていれば、この最尤系統樹を初期系統樹として再びTBRによる変形をほどこす。この過程を尤度が改善されなくなるまで繰り返す。こうして得られた最終的な系統樹を最尤系統樹とする。OTUが20程度の分子系統樹であれば、この方法により最尤系統樹が得られることが経験的に示されている(Katoh & Miyata, in press)。

系統樹の確からしさは、系統樹上の各分岐点において、分岐点の局所ブートストラップ確率(LBP, Adachi & Hasegawa, 1996)を求めて検証した。このLBPの値は、分岐点の内部枝について局所的再配置を施し、得られた三種類の樹形を比較した際に、もとの樹形が得られるブートストラップ確率である。したがってLBPの値は、本来のブートストラップ確率

の値よりも大きくなることに注意する必要がある。

C. 近隣結合法

OTUが30以上の場合には、近隣結合法(Saitou & Nei, 1987)を用いた。近隣結合法は、個々の配列間のアミノ酸置換数から得られる距離行列をもとに、最小進化の原理により分子系統樹を推定する。この方法は、計算が簡単で計算量も少ないことから、広く使われている方法である。ただし、分子進化速度が系統によって異なったり、多重置換効果が大きくなる場合には（すなわち距離行列が正しく推定できない場合には）、誤った系統樹を推定することが知られている(Hasegawa & Fujiwara, 1993)。本論文ではOTUが30以上となる場合、また発見的探索による最尤法を行う際の初期系統樹を得る場合に限って近隣結合法を用いた。近隣結合法では距離行列を正しく推定することが重要である(Hasegawa & Fujiwara, 1993)。そこで、距離行列の推定には最尤法を用いた。すなわち、最尤法を二本のOTUからなる系統樹に適用することで、二本の配列間のアミノ酸置換数を推定した。アミノ酸置換の確率法則としてはJTTモデル(Jones et al., 1992)を用いた。

系統樹の確からしさは、ブートストラップ検定(Felsenstein, 1985)により、系統樹上の各分岐点において、分岐点以下のOTUが単系統となるブートストラップ確率を求めて検証した。

遺伝子重複が起きた時期の推定

分子系統樹上の分岐点は、生物種の分岐もしくは遺伝子重複を示す。ある分岐点が生物種の分岐であるか遺伝子重複であるかの判断は、その分岐が生物の系統関係を反映しているか否かに基づいて行う。したがって、遺伝子重複の推定をする際には、生物の系統関係を正しく仮定する必要がある。本論文においては、形態的、分子系統学的に支持される図2の系統関係を仮定した。

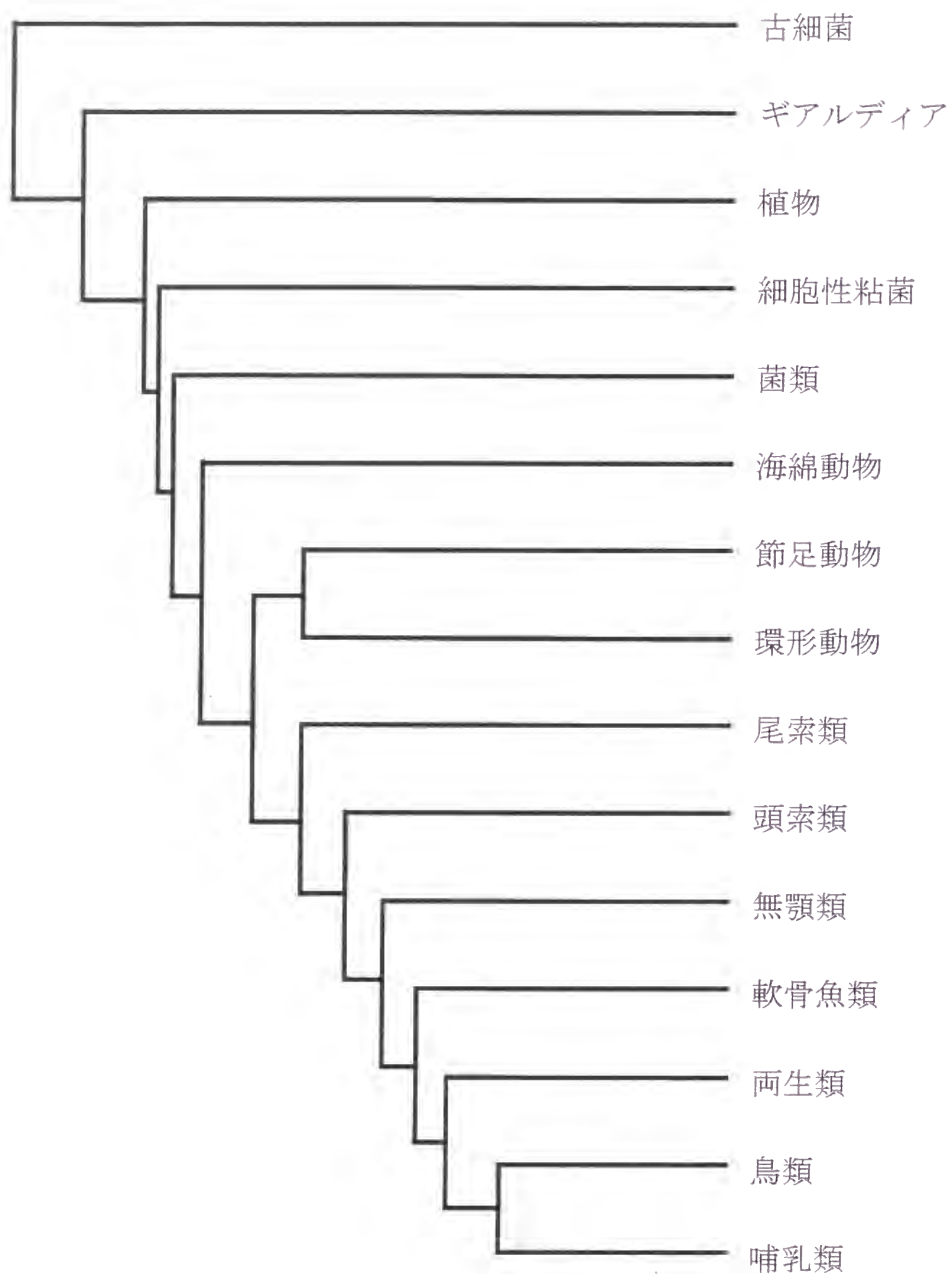


図2 本論文において仮定した生物の系統関係

ギアルディアの分岐はSogin et al., 1989; Hashimoto et al., 1994、植物及び菌類の分岐はNikoh et al., 1994、細胞性粘菌の分岐はKuma et al., 1995b、海綿動物の分岐はWillmer, 1990; Wainright et al., 1993; Kobayashi et al., 1993、尾索類及び頭索類の分岐はWillmer 1990; Wada and Satoh, 1994に基づいた。それ以外の系統関係はRomer & Parsons, 1986を参考にに基づいた。

遺伝子重複が起きた時期及びその回数の推定は、分子系統樹に基づき行った。時期の推定は生物種の分岐を指標に行った。ただし分子系統樹中には、分岐点のブートストラップ確率が低く、信頼性の低い分岐も含まれる。そこで、遺伝子重複が起きた時期及び回数の信頼性を統計的に検定する目的で、ブートストラップ検定を行った。

A. 最尤法で分子系統樹を推定した場合

最尤法で尤度を計算した全ての系統樹のうち、最尤系統樹との尤度差が一標準誤差範囲内にあり棄却されない系統樹の中には、真の系統樹が含まれる可能性がある。そこで、一標準誤差範囲内で棄却されない系統樹全てについて、遺伝子重複が起きた回数を数え上げた。そして、各系統樹について得られた遺伝子重複回数を、各系統樹のブートストラップ確率を規格化した値で重みづけし、足し合わせた。この値を、ブートストラップ検定により得られた遺伝子重複回数とした。すなわち、 b_x を樹形 x のブートストラップ確率、 B_x を樹形 x の規格化されたブートストラップ確率、 N_x を樹形 x に基づいて得られた遺伝子重複の回数とすれば、ブートストラップ検定による遺伝子重複回数 N は、以下の式で求められる。

$$B_x = b_x / \sum_{i=1}^n b_i$$

$$N = \sum_{i=1}^n N_i B_i \quad (n \text{ は一標準誤差範囲で棄却されない系統樹の総数})$$

B. 近隣結合法で系統樹を推定した場合

ブートストラップ検定(Felsenstein, 1985)により分岐点のブートストラップ確率を得る方法と同じ手順で100の系統樹を生成し、それぞれの系統樹について遺伝子重複回数を数え上げ、平均及び分散を求めた。

分子進化速度の推定

分子系統樹の枝の長さは、その系統で蓄積したアミノ酸置換を反映している。したがって分子系統樹の枝の長さに、年代の情報を与えることで、各系統における分子進化速度の推定が可能である。分子進化速度を推定する際には、表5の生物の分岐年代を仮定した。また、系統樹上異なる系統間で分子進化速度の平均を求める場合、すなわち複数の枝の平均長を求める場合には、各系統に属する遺伝子数に影響されないよう考慮する必要がある。そこで、複数の枝の平均分子進化速度は、分岐点から出る2本の枝長の平均を順々に足し合わせることで求めた。例えば図3において、黒丸からA～Eに至る系統までの平均分子進化速度は以下の式から求めた。

$$t_{average} = [(t_1+t_2)/2+t_3 + \{ (t_4+t_5)/2+t_6+t_7 \} / 2+t_8] / 2$$

表5 本論文において仮定した分岐年代

分岐	分岐年代（億年前）
動物/植物/菌類	10.7
海綿動物/真正後生動物*	9.4
節足動物/脊椎動物	7.0
軟骨魚類/その他の有顎類	4.0

Dickerson, 1971; Dayhoff, 1978; Nikoh et al., 1997に基づいた。*真正後生動物とは、海綿動物以外の多細胞動物の総称である。

同義置換数及び非同義置換数の推定

系統樹推定はアミノ酸配列に基づいて行ったが、配列間の分岐が比較的最近に起きた場合には、必要に応じて塩基配列の比較も行った。塩基配列の同義置換（アミノ酸置換を伴わない置換）数及び非同義置換（アミノ酸置換を伴う置換）数は、Miyata & Yasunaga (1980)の方法により推定した。多重置換の補正はJukes & Cantor(1969)の方法で行った。

PEST配列の検索

PEST配列とは、寿命の短いタンパク質に共通にみられるモチーフである(Rogers et al., 1986)。単離したPTP遺伝子にこのPEST配列が存在するか調べる目的で、PEST配列検索ソフトウェアPESTfind ([http://embl.bcc.univie.ac.at/embnet/tools/bio/PEST/ find](http://embl.bcc.univie.ac.at/embnet/tools/bio/PEST/find))を用いた。

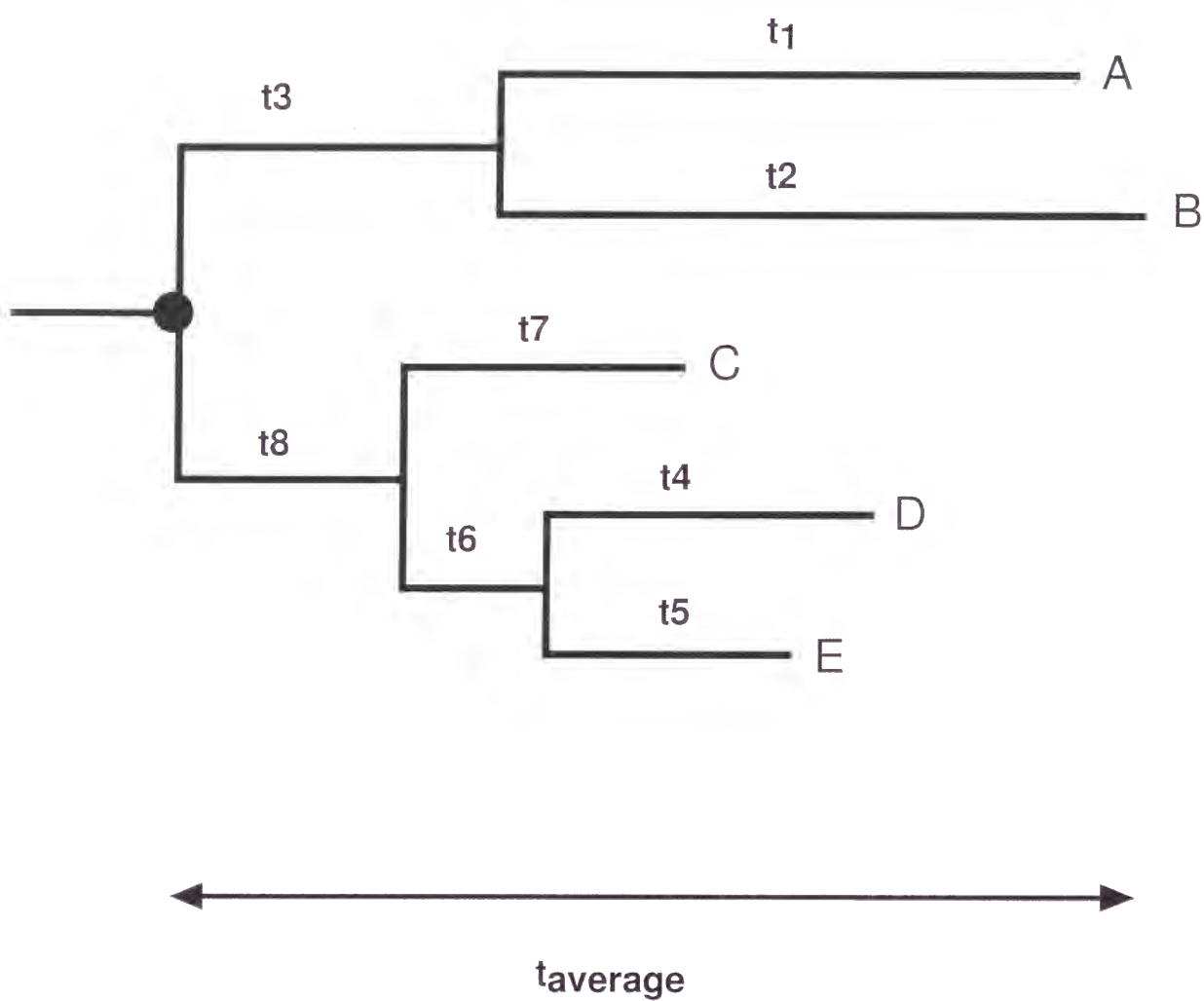


図3 分子進化速度の推定方法
 $t_1 \sim t_8$ は枝の長さを示す。

結果と考察

1. プロテインチロシンホスファターゼ(PTP)遺伝子族

1-1. PTP遺伝子族のサブタイプの多様化

動物には数多くのPTP遺伝子が存在することが知られている。その多くはPTPドメイン以外に多様なドメインをもち、大きく二つ、受容体型と細胞質型とに分類される（図4）。これに対して植物・細胞性粘菌・菌類のPTP遺伝子は全て細胞質型で、PTPドメイン以外のドメインをもたない。このことから、動物で特異的に見られる受容体型PTP遺伝子や多様なドメインをもつPTP遺伝子は、動物の系統で独自に作り出されたものと考えられる。動物に見られるPTP遺伝子族の多様なドメインは、結合するリガンドや基質特異性、細胞内局在の決定等に関わっている(Mauro & Dixon, 1994 for review)。したがって、動物は多様なドメインをもつPTP遺伝子を作り出し、各遺伝子は多様かつ複雑な機能を担うよう進化した、と考えることができる。同じドメイン構造を持つ遺伝子群は、共通の機能を担うまとまったグループと考えられることから、本論文ではこれをサブタイプと呼ぶことにする。現在、動物には10種類以上のサブタイプが知られており、それぞれが固有の機能を果たしている(Charbonneau & Tonks, 1992; Hunter, 1995; Streuli, 1996 for review)。

動物にみられるPTP遺伝子の多様なサブタイプは、遺伝子重複とそれに続くドメインシャッフリングにより作り出されたものと考えられる。それでは、動物の進化の過程において、PTP遺伝子族の多様なサブタイプが作り出された時期はいつだろうか。この問題を明らかにする目的で、動物の中で最初に分岐した(Willmer, 1990; Wainright et al., 1993; Kobayashi et al., 1993)原始的な多細胞動物である海綿動物に注目した（図5）。海綿動物は、細胞同士の接着が弱く細胞間の協調もほとんどみられない原始的な多細胞動物である(Margulis & Schwartz, 1998)。この海綿動物が他の動物と分岐する時点では、どれ程のサブタイプが作り出されていたのだろうか。いいかえれば、海綿動物にはどれだけのサブタイプが存在するのだろうか。また、動物の多細胞性の進化を分子レベルから理解する上で、受容体型のPTP遺伝子は重要である。受容体型PTP遺伝子は細胞接着に深く関与しており、多細胞動物の進化に重要な役割を果たしたと考えられる。例えばPTP μ 遺伝子はホモフィリックな結合をすることが知られており、細胞接着の制御に関与することが示唆されている(Sap et al., 1994)。それでは、原始的な多細胞動物である海綿動物に、この受容体型PTP

遺伝子はどれだけ存在するのだろうか。以上のことを明らかにする目的で、RT-PCRによりカワカイメンからPTP遺伝子を探索し、分子進化学的解析を行った。

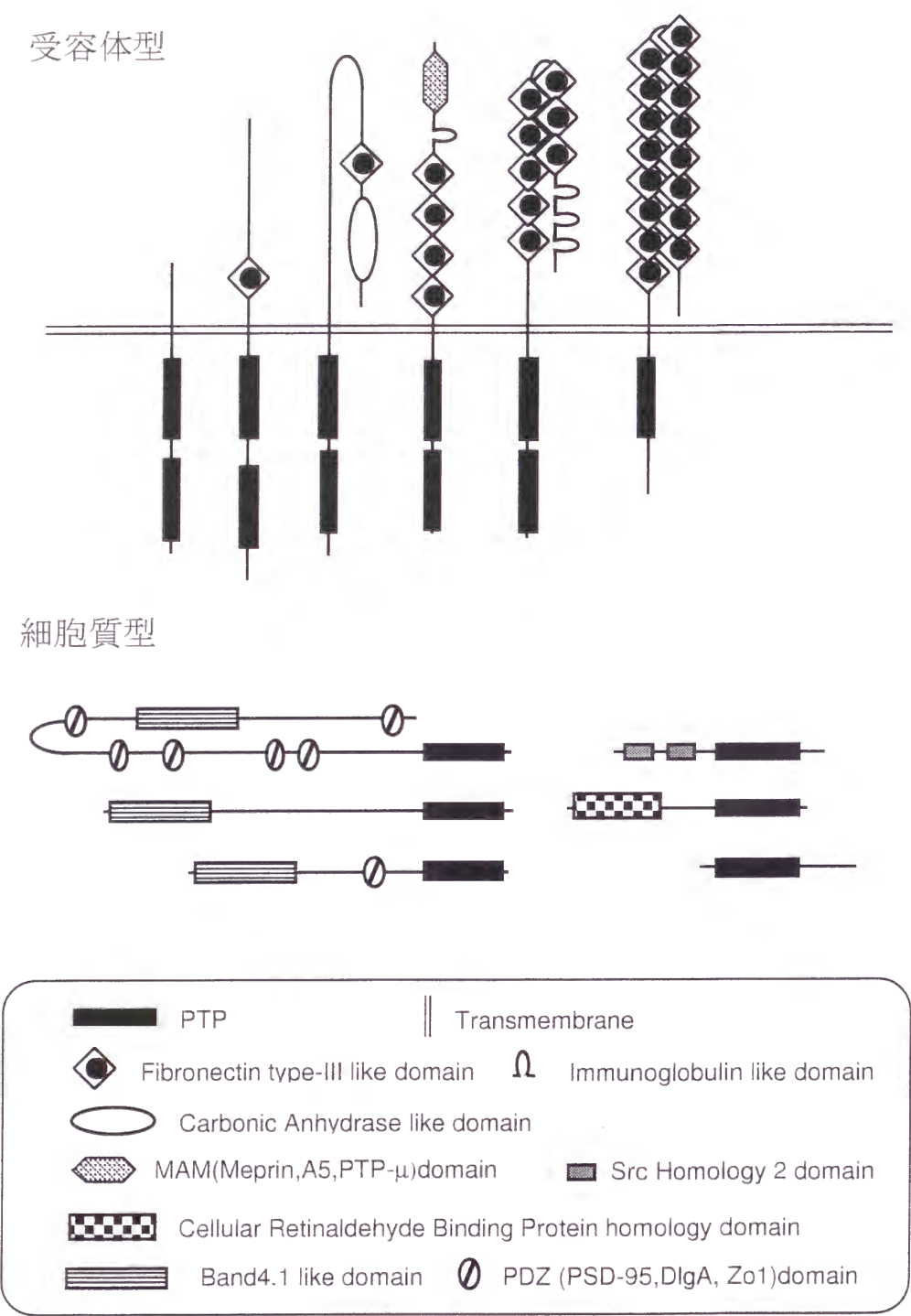


図4 動物のPTP遺伝子の多様なサブタイプ

多様なドメインをもつ動物のPTP遺伝子を模式的に示した。ドメインはStreuli, 1996; Stoker, 1996に基づいた。本論文において「ドメイン」とは、既知の配列と相同性が認められ、何らかの機能を果たしていると考えられる遺伝子内の領域を指す。すなわち、進化の過程でエクソンシャッフリングなどの機構により、PTPドメインと連結されたと考えられる遺伝子上の単位である。

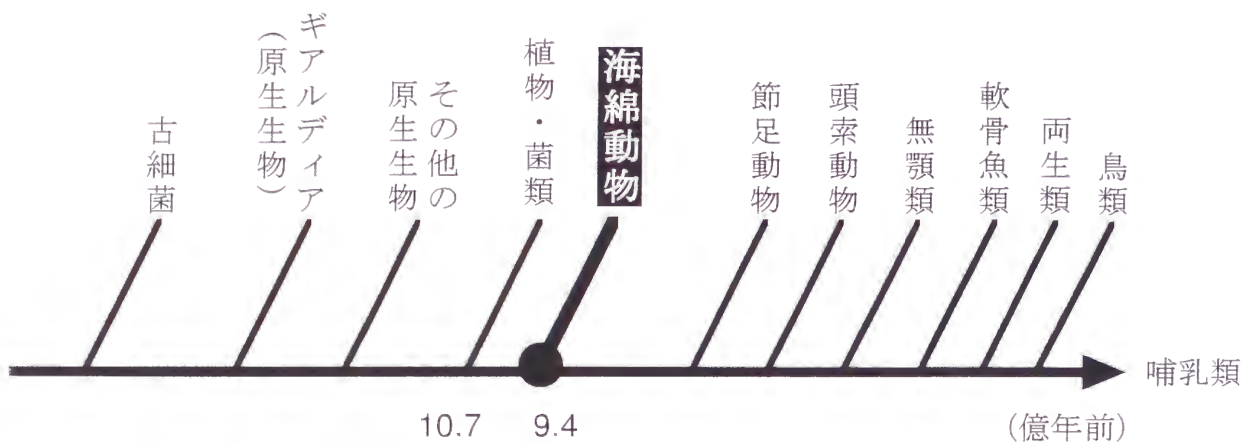


図5 海綿動物の系統的位置

本節で注目する海綿動物の系統的位置を模式的に示した。枝の長さは概ね時間を反映しているが正確ではない。数字は分岐年代を示す(Nikoh et al., 1997)。

結果

カワカイメンのPTP遺伝子のクローニング

RT-PCRによりカワカイメンからPTP遺伝子の探査を試みた。その結果、カワカイメンより10種類のPTP遺伝子の単離に成功した（塩基配列及び推定アミノ酸配列は付録2-1に示した）。塩基配列を決定した領域は、PTPドメインのN末側約1/3から終止コドンまでに相当する。

単離したカワカイメンのPTP遺伝子は、菌類のPTP遺伝子と26～41%、他の動物のPTP遺伝子と30～71%の一致度を示し、PTP signature motif 'HCXXGXXR(S/T)' (Denu et al., 1996)が保存していた（図6）。したがって、これらの遺伝子はPTPとして機能しているものと思われる。

単離したカワカイメンの10種類のPTP遺伝子のうち、3種類はPTPドメインを二つ持っていた（図6）。現在までに知られている動物の細胞質型PTP遺伝子は全て、PTPドメインを一つしか持たない。これに対し、受容体型PTP遺伝子の多くは、PTPドメインを二つ繰り返して持つ。したがって、この3遺伝子は受容体型のPTP遺伝子であることが予想される。

human	PTPN1 (M33689)	(208)	GPVVV HCSAGIGRS GTFLADTCLLLMDKRKD	(195)
1) PTP domain				
sponge	sPTPR3 (AB019128)	(107)	VPLL VHCSAGVGR TGTFIALDTLLDRVRSETS	(44)
sponge	sPTPR7 (AB019131)	(107)	GPTVV HCSAGVGR SGAFIAIHMGITEFKRMKI	(40)
sponge	sPTPN1 (AB019129)	(110)	PPVL VHCSAGVGR SGAFCAIDYCIDELRDLGQ	(38)
sponge	sPTPN2 (AB019130)	(101)	HPIL AHCSAGLGR TGVLVAVHTALEYHKAGHK	(704)
sponge	sPTPN3	(105)	GPCIV HCSAGIGRS GTFLVDVCLKKIESTRD	(337)
sponge	sPTPN6 (AB019132)	(115)	GPIVV HCSAGIGRT GTFFIVIDVIQDLINHQGW	(82)
sponge	sPTPN8 (AB019133)	(107)	VPVL VHCSAGVGR SGVFVLMDFLMAKVDCGDE	(41)
2) first PTP domain				
sponge	sPTPR2B (AB019126)	(107)	GPM AVHCSAGVGR TGTFICIDYVLKQIEGEAI	(339)
sponge	sPTPR4 (AB019125)	(102)	GPM VHCSAGVGR TGTFITLHSQMNRLREERN	(337)
sponge	sPTPR5 (AB019127)	(101)	GPLL VHCSAGVGR TGTVIAIDYCMKQIHMEGE	(313)
3) second PTP domain				
sponge	sPTPR2B (AB019126)	(401)	GPIT VHCNDGIGRT GTFCAYSMMDRVKVEQV	(45)
sponge	sPTPR4 (AB019125)	(394)	KPIV VHCSAGCGRT GTYYVAISILLERLRTEGA	(45)
sponge	sPTPR5 (AB019127)	(371)	NAIL VHDTTSLG PSGIFCALWHCMERLRLERV	(43)
			* * *	

図6 カワカイメンのPTP遺伝子の活性中心付近のアライメント

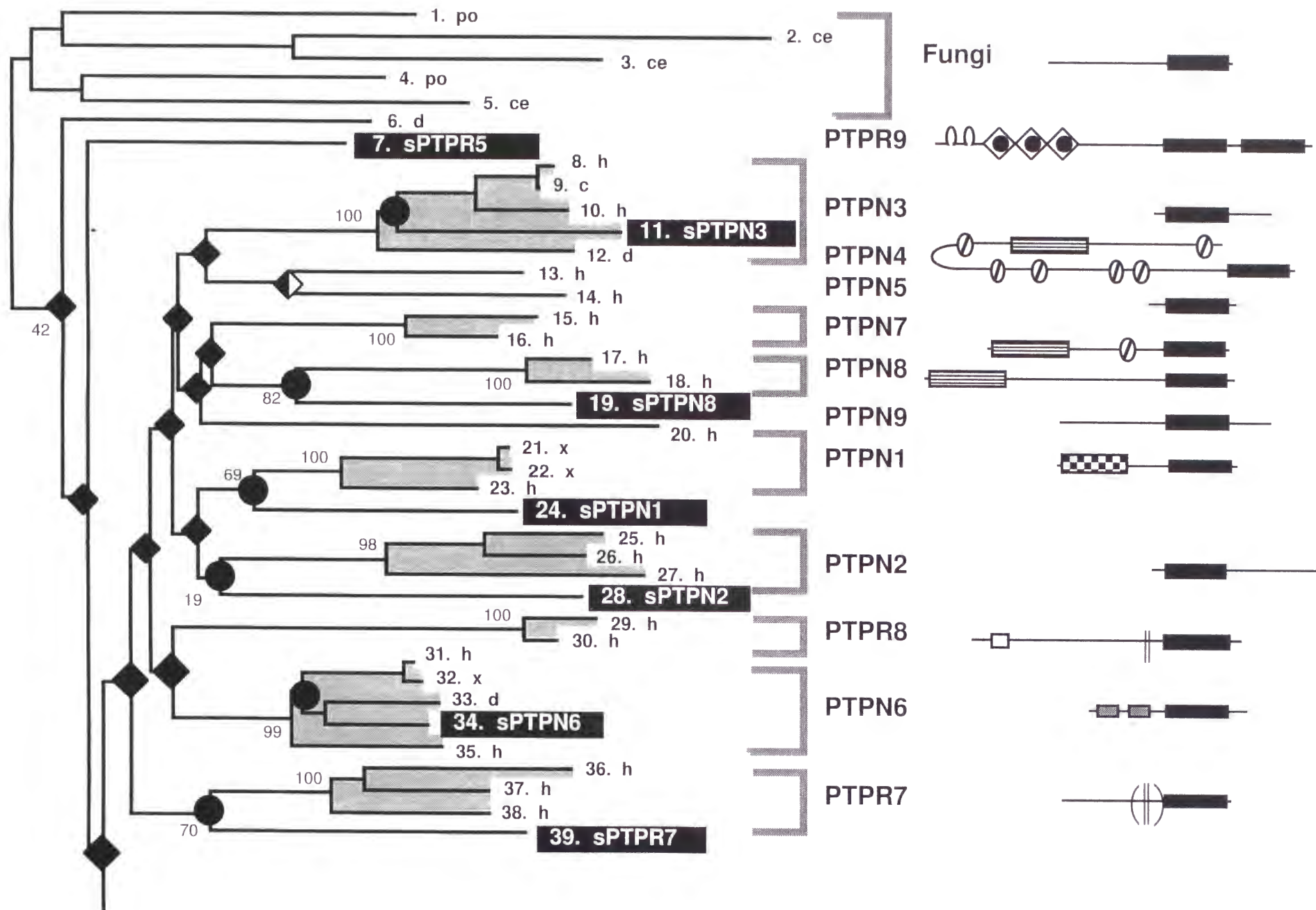
単離されたカワカイメンのPTP遺伝子におけるPTPドメインの活性中心付近のアライメントを示した。PTP signature motif 'HCXXGX_{XX}R(S/T)'(Denu et al., 1996)を白抜き文字で示した。1)PTPドメインを一つもつかワカイメンの遺伝子のPTPドメイン。2)PTPドメインを二つもつかワカイメンの遺伝子のN末側のPTPドメイン。3)PTPドメインを二つもつかワカイメンの遺伝子のC末側のPTPドメイン。比較のためヒトのPTPN1遺伝子を最上に示した。遺伝子名は図7の分子系統解析に基づき名付けた。括弧内はその領域に存在するアミノ酸の総数。

PTP遺伝子族の分子系統樹の推定

PTP遺伝子族の多様なサブタイプが遺伝子重複により作り出された時期を知る目的で、既存のデータベース中に存在するPTP遺伝子と、単離されたカワカイメンのPTP遺伝子を併せて、近隣結合法により分子系統樹を推定した。その結果、動物のPTP遺伝子群は、植物・細胞性粘菌・菌類のPTP遺伝子群に対して単系統群を形成した。このことは、動物に存在する多様なサブタイプが、植物・細胞性粘菌・菌類との分岐後に、動物の系統で独自に作り出されたことを意味する。この結果は、動物のみが多様なドメインをもつということとも一致する。図7に、アウトグループとして動物に最も近縁な(Nikoh et al., 1994)菌類のPTP遺伝子群のみを用いた、より信頼性の高い分子系統樹を示す。

図7の分子系統樹から、同じサブタイプに属する（同じドメイン構造をもつ）遺伝子は単系統群を形成することがわかった（網掛けで示した）。この分子系統樹をもとに、PTP遺伝子族は18種類のサブタイプに分類された。サブタイプの名前は Charbonneau and Tonks (1992)及び Brady-Kainay and Tonks (1995)を参考に、受容体型のものをPTPR1～9、細胞質型のものをPTPN1～9と名付け、系統樹の右側に示した。

単離した10種類のカワカイメンの遺伝子（白抜き文字で示した）のうち、8種類はいずれかのサブタイプと明らかに近縁だった。すなわち、sPTPR2B、sPTPR3、sPTPR7、sPTPN1、sPTPN2、sPTPN3、sPTPN6、sPTPN8は、それぞれPTPR2B、PTPR3、PTPR7、PTPN1、PTPN2、PTPN3、PTPN6、PTPN8サブタイプのホモログだと考えられた。残りの2種類の遺伝子、sPTPR4とsPTPR5は、明らかにどのサブタイプに属するとはいえなかった。ただし、アウトグループの菌類の遺伝子を除き、分子系統樹推定に用いるアミノ酸領域を増やした分子系統樹からは、sPTPR4はPTPR4サブタイプに、sPTPR5はPTPR5サブタイプに属することが示されている。特に、sPTPR5にはPTPR5サブタイプに認められる特徴が存在していた。すなわち、二つめのPTPドメインの活性中心のシステイン残基がアスパラギン酸残基に変化している、という特徴がみられた（図6）。このことは、sPTPR5がPTPR5サブタイプに属することを支持する。



0.1
substitutions/site

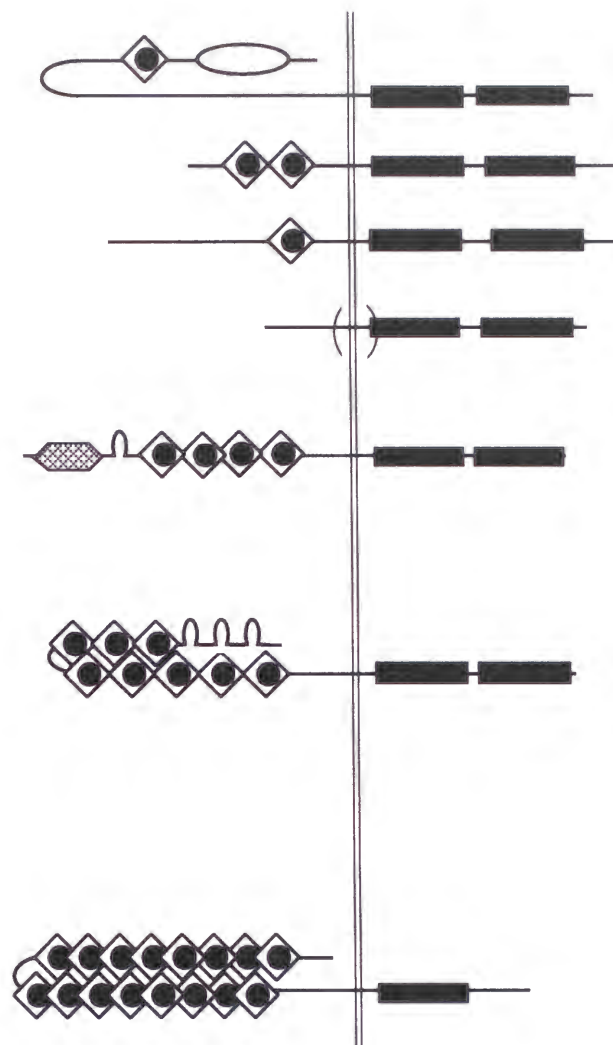
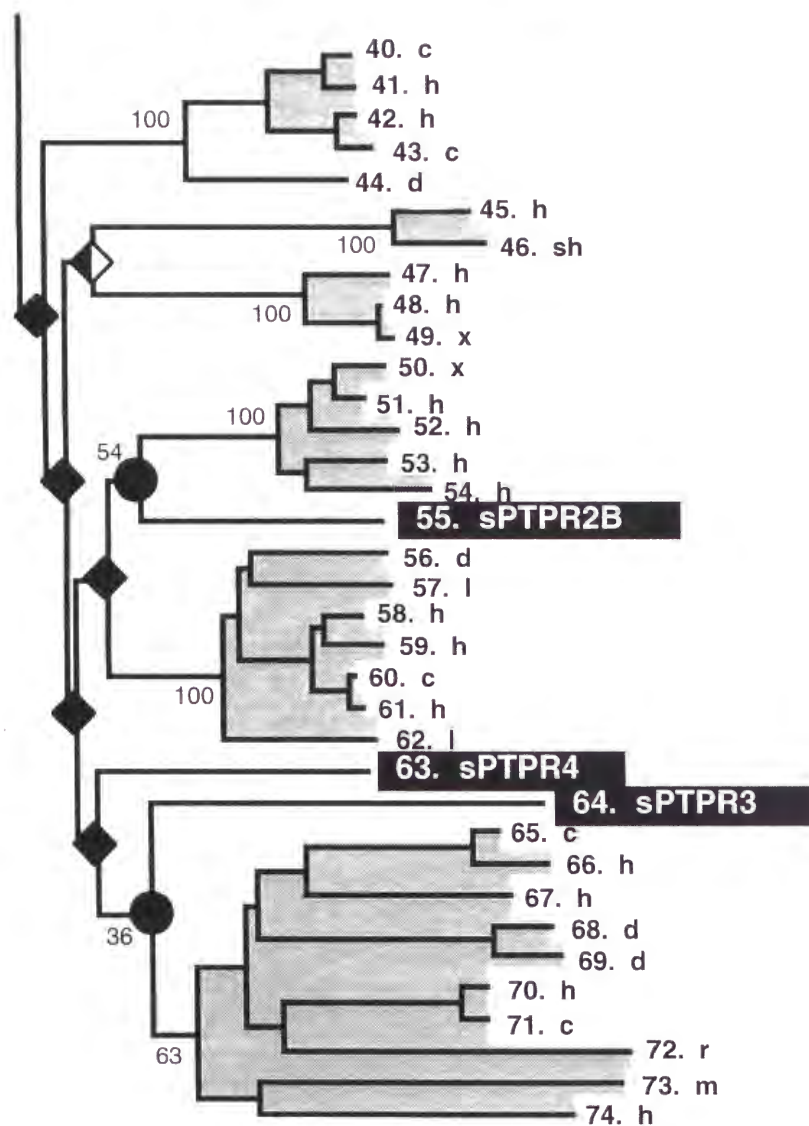


図7 PTP遺伝子族の分子系統樹

PTPドメイン約3分の2に相当するアライメント領域（133アミノ酸座位）に基づき、近隣結合法により推定した分子系統樹を示した。アライメントは付録5-1に示した。アウトグループには菌類のPTP遺伝子を用いた。遺伝子配列を決定したカワカイメンの遺伝子は白抜き文字で示した。各系統の記号は以下を表す。h, human; m, mouse; r, rat; c, chicken; x, *Xenopus*; sh, shark; d, *Drosophila*; l, leech; ce, *Saccharomyces cerevisiae*; po, *Schizosaccharomyces pombe*. 系統樹推定に用いた遺伝子の名前及びGenBankのアクセッションナンバーは表9に示した。同じドメイン構造を保持する遺伝子群を網掛けで示し、同じサブタイプに分類した。サブタイプ名はCharbonneau and Tonks (1992)及び Brady-Kainay and Tonks (1995)を参考に右側に記した。ドメインの模式図は図4と同じである。本論文における「ドメイン構造」とは、ドメインの種類及びその並び順を指し、ドメインの繰り返しの数は問わない。なぜなら繰り返しの数は、遺伝子内における遺伝子変換などの機構により変化する可能性があるからである。図に示したドメイン構造は、サブタイプ内のヒトの遺伝子で最長のものを代表させて表した。また膜貫通部分の括弧は、サブタイプ内の遺伝子により、あるいは選択的スプライシングにより、受容体型の遺伝子と細胞質型の遺伝子の両方が存在することを意味する。ノード付近に示した数字はブートストラップ確率を示す。

黒丸は海綿動物と真正後生動物の分岐を、菱形は異なるサブタイプを作り出した遺伝子重複を示す。このうち黒い菱形は菌類と動物の分岐後、海綿動物と真正後生動物の分岐前に起きたものを、白黒の菱形は、菌類と動物の分岐後、海綿動物と真正後生動物の分岐前か後か不明のものである。

尚、PTPドメインを二つ繰り返してもつ遺伝子については、両ドメインを含めた分子系統解析から、その繰り返し構造が、サブタイプの分岐前に起きた1回のイベントで作られ出したことが示された。遺伝子重複の起きた時期を推定する際に両ドメインを含めた分子系統樹を用いると、同じ遺伝子重複を2度数え上げることになる。したがって、PTPドメインを二つもつ遺伝子に関しては、N末端側のPTPドメインを分子系統樹推定に用いた。N末端側のPTPドメインを用いたのは、C末端側PTPドメインの中には酵素活性をもたず進化速度が高い遺伝子が存在するため、N末端側のPTPドメインの方が系統樹推定に適当だと考えたからである。(Schaapveld et al., 1997)。

PTP遺伝子族のサブタイプを作り出す遺伝子重複が起きた時期の推定

図7の分子系統樹から、PTP遺伝子の多様なサブタイプが遺伝子重複により作り出された時期が推定できる。分子系統樹上で、異なるサブタイプを作り出した遺伝子重複に、菱形の印を付した。すると、そのような遺伝子重複19回のうち17回が、動物界最古の種分岐である海綿動物とその他の動物（真正後生動物）の分岐（黒丸）前に起きたことがわかった（黒い菱形）。残りの2回の遺伝子重複については、遺伝子重複が起きた時期が、海綿動物と真正後生動物の分岐の前なのか後なのかは明らかでない（白黒の菱形）。また、海綿動物と真正後生動物の分岐後には、異なるサブタイプを作り出した遺伝子重複はみられなかった。

この結果は、図7の分子系統樹に基づいて得られた結果である。図7の分子系統樹には、分岐点のブートストラップ確率が低く、信頼性の低い分岐も含まれる。そこで、異なるサブタイプを作り出した遺伝子重複が起きた時期を統計的に検定する目的で、ブートストラップ検定を行った（図8）。その結果、異なるサブタイプを作り出した遺伝子重複のうち、菌類と動物の分岐後、海綿動物と真正後生動物の分岐前（これをI期とする）に起きたものは平均で14.5回、海綿動物と真正後生動物の分岐後（これをII期とする）に起きたものは平均で1.1回、I期なのかII期なのか不明なものが平均で2.8回となった。この値は図7の分子系統樹に基づいて得られた値と概ね一致する。このことから、ほとんどのサブタイプがI期に集中的に作り出されたことが確かめられた。

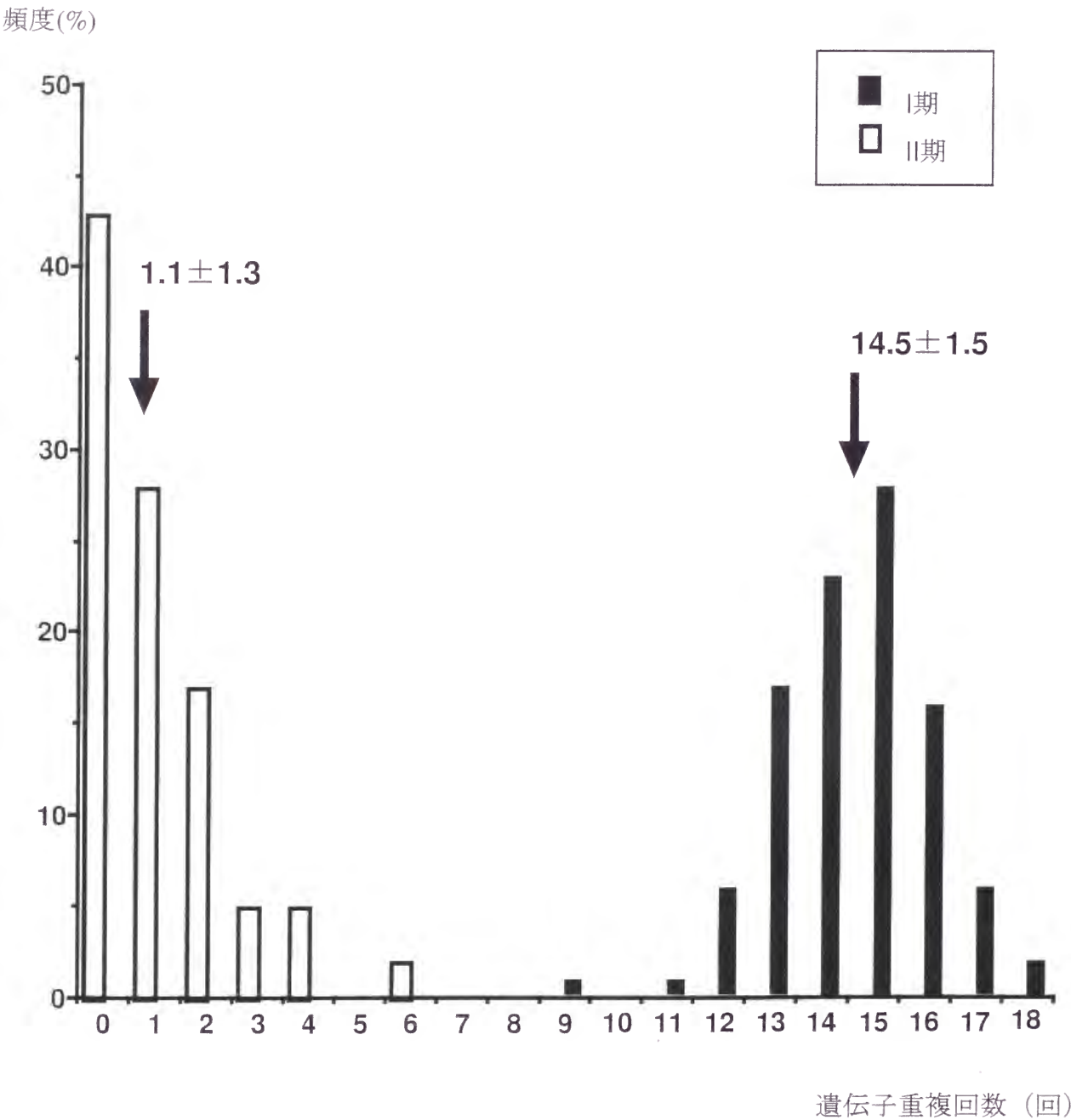


図8 PTP遺伝子族のサブタイプを作り出す遺伝子重複が起きた時期のブートストラップ検定

ブートストラップ検定に基づき、異なるサブタイプを作り出す遺伝子重複が起きた時期及び回数の頻度分布を求め、ヒストグラムに示した。矢印は平均を示す。異なるサブタイプを作り出す遺伝子重複のうち、I期に起きたものを黒、II期に起きたものを白で表した。

PTP遺伝子族のサブタイプを作り出すドメインシャッフリングが起きた時期の推定

PTP遺伝子の多様なサブタイプを作り出したドメインシャッフリングが起きた時期についてはどうだろうか。PTPR2A、PTPR3、PTPN6サブタイプに関しては、これらのサブタイプに属する節足動物と脊椎動物の遺伝子が存在する。すなわち節足動物と脊椎動物の間でドメイン構造が保存していることから、少なくとも節足動物と脊椎動物の分岐以前に、ドメインシャッフリングが起きたと考えられる。単離したカワカイメンのPTP遺伝子が、ホモログと思われる高等動物の遺伝子と同様のドメイン構造を保持していた場合には、ドメインシャッフリングが起きた時期はさらに古くに遡ることになる。すなわち、遺伝子重複のみでなくドメインシャッフリングもまたI期に起き、サブタイプの特徴であるドメイン構造がI期には完成していたことになる。PTPN2サブタイプとPTPN3サブタイプについては、この問題に関する示唆が得られる。なぜなら、ドメイン構造は存在しないものの、これらのサブタイプのC末端側には特徴ある配列が存在しているからである。PTPN2サブタイプはPTPドメインに続き、そのC末側にPEST配列と呼ばれる寿命の短いタンパク質に共通のモチーフが存在している(Rogers et al., 1986; Matthews et al., 1992)。カイメンのsPTPN2遺伝子のC末側の領域にもPEST配列が存在することが、PEST配列検索ソフトウェアPESTfind(材料と方法の章参照)の結果から示された(付録2-1参照)。また、PTPN3サブタイプはPTPドメインのC末側に100~200アミノ酸程の保存されていない領域があり、そのC末端に疎水性のアミノ酸が30アミノ酸前後連続している。この疎水性のアミノ酸の連続は、小胞体などの細胞内膜系との結合に寄与することが示唆されている(Mauro & Dixon, 1994)。カワカイメンのsPTPN3遺伝子にもまた、PTPドメインのC末側に250アミノ酸程に続き、C末端に疎水性のアミノ酸が25残基程連続していた(付録2-1参照)。したがって、PTPN2サブタイプ及びPTPN3サブタイプにおいては、機能に関して重要な役割をもつと考えられる特徴的な配列が、カワカイメンのPTP遺伝子にも保存されていることがわかった。このことから、ドメインシャッフリングもまたI期に起きたことが示唆される。

ただし、PTPR5サブタイプにおいては、脊椎動物に見られるCAドメイン(Carbonic Anhydrase like domain)が節足動物では欠けている。したがってPTPR5サブタイプについては、ドメインシャッフリングが起きた時期が、脊椎動物と節足動物の分岐後である可能性がある。あるいはI期にドメインシャッフリングが起きた後、節足動物の系統でCAドメインが失われたのかもしれない。

分子進化速度の比較

図7の分子系統樹はI期にアミノ酸置換の頻度（分子進化速度）が高いことも示している。分子系統樹の枝の長さは、その系統で蓄積したアミノ酸置換を反映している。したがって分子系統樹の枝の長さに、年代の情報を与えることで、各系統における分子進化速度の推定が可能である。そこで、菌類と動物の分岐年代を10.7億年前、海綿動物と真正後生動物の分岐年代を9.4億年前として(Nikoh et al., 1997)、各サブタイプについて、I期とII期における分子進化速度の比較を行った。その結果、サブタイプごとに分子進化速度は異なるものの、総じてI期はII期に比べて分子進化速度が高く、平均で約6倍となることがわかった（表6）。すなわち、I期には遺伝子重複のみでなくアミノ酸の置換も多く起きていることがわかった。

表6 分子進化速度の比較

サブタイプ名	V _I	V _{II}	V _I /V _{II}
PTPR2B	1.5 x10 ⁻⁹	0.30 x10 ⁻⁹	5.0
PTPR3	1.6	0.48	3.3
PTPR7	1.8	0.53	3.4
PTPN1	2.4	0.43	5.6
PTPN2	1.9	0.65	3.0
PTPN3	3.9	0.35	11.0
PTPN6	2.8	0.24	11.6
PTPN8	2.9	0.51	5.6

図7の分子系統樹に基づき、1アミノ酸座位当たり、1年当たりに起きたアミノ酸の置換数（進化速度）を推定した。V_IはI期の分子進化速度を、V_{II}はII期の分子進化速度を表す。各サブタイプの最も古い分岐を海綿動物と真正後生動物の分岐（9.4億年前）と仮定して計算した。この仮定はV_Iを過小評価することになり、本論文の結論に対して不利に働く。

考察

結果のまとめ

RT-PCRによりカワカイメンから計10種類のPTP遺伝子の単離に成功した。

単離した遺伝子を含めて分子進化学的解析を行った結果、動物のPTP遺伝子の多様なサブタイプは、菌類と動物の分岐（約10.7億年前、Nikoh et al., 1997）後、海綿動物と真正後生動物の分岐（約9.4億年前、Nikoh et al., 1997）前の、約1億年の間に、急激に作り出されたことが示された（図9）。また、海綿動物と真正後生動物の分岐後現在に至るまでの、約9億年の間には、新しいサブタイプを作り出す遺伝子重複は見られないこともわかった。

原始的な多細胞動物であるカワカイメンからも、細胞接着に重要な役割を果たすPTPμのホモログを含む、4種類の受容体型と予想されるPTP遺伝子が存在した。またこの他の受容体型サブタイプも、その起源が非常に古いことから、カワカイメンにも存在することが示唆される。

多様なサブタイプを作り出した遺伝子重複が集中的に起きた時期は、そうでない時期と比べて、分子進化速度が高かった。

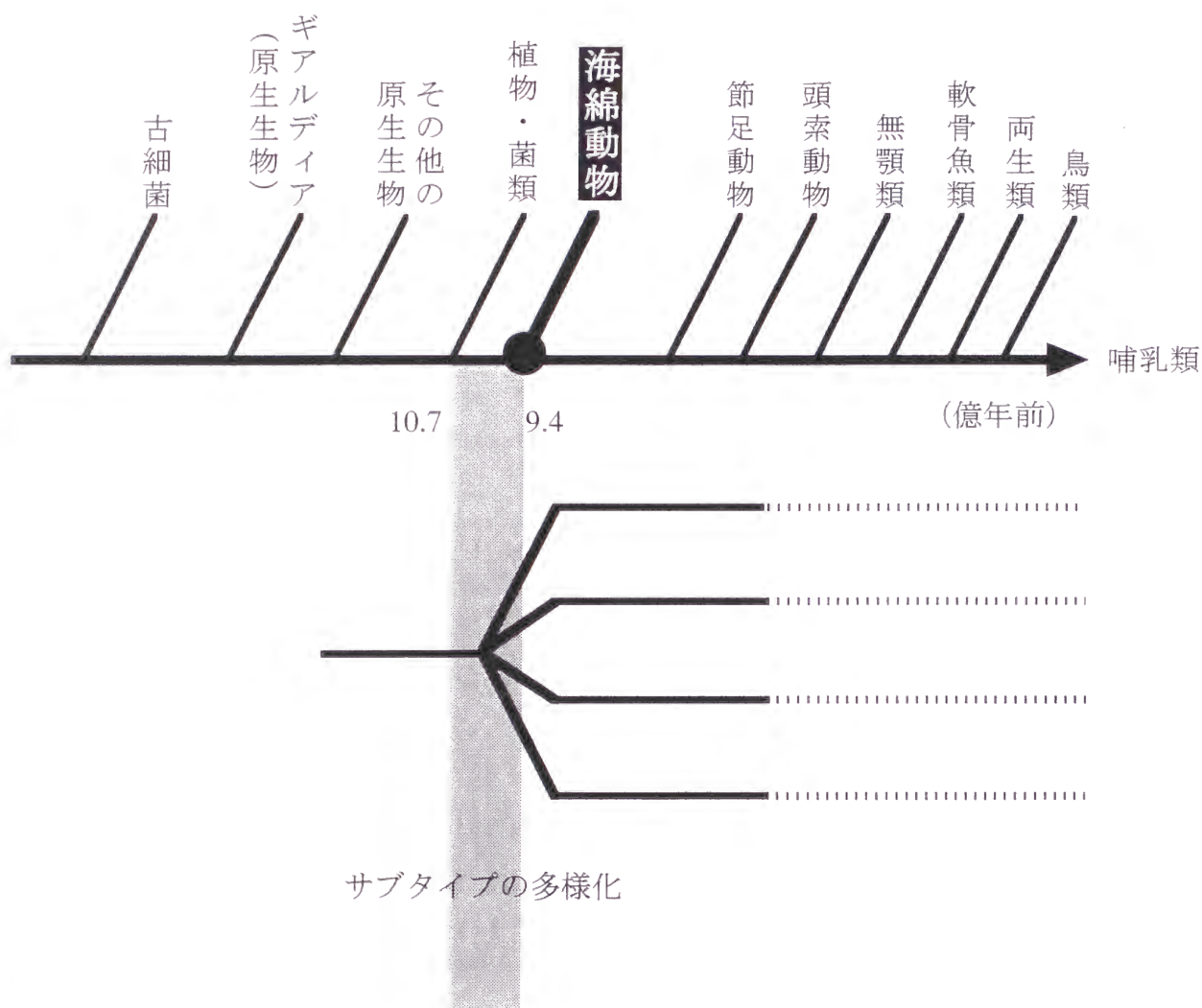


図9 PTP遺伝子族のサブタイプの多様化のまとめ

PTP遺伝子族のサブタイプの多様化の時期を模式的に示した。本論文の結果から明らかとなった、サブタイプの多様化が見られた時期を網掛けで示した。枝の長さは概ね時間を反映しているが正確ではない。数字は分岐年代を示す(Nikoh et al., 1997)。

他の遺伝子族との比較

PTP遺伝子族の多様なサブタイプが、遺伝子重複により作り出された時期は、菌類と動物の分岐後、海綿動物と真正後生動物の分岐前に集中していた。同様の結果は、細胞のシグナル伝達に関わる他の遺伝子族、Gタンパク質 α サブユニット($G\alpha$)遺伝子族、ホスホジエステラーゼ(PDE)遺伝子族、ホスホリパーゼC (PLC)遺伝子族・プロテインキナーゼC (PKC)遺伝子族や、形態形成に関与する転写因子である *Pax* 遺伝子族でもみられる（表7、Suga et al., 1999a; 1999b; Koyanagi et al., 1998a; 1998b; Hoshiyama et al., 1998）。また小柳らにより、PDE遺伝子族においては、多様なサブタイプを作り出したドメインシャッフリングが起きた時期もまた、海綿動物と真正後生動物の分岐前であることが示されている (Koyanagi et al., 1998a)。

表7 多様なサブタイプを作り出す遺伝子重複が起きた時期及びその回数

遺伝子族	I期	II期
$G\alpha$	8.5 ± 0.9 (回)	1.1 ± 1.2 (回)
PTK	22.5 ± 2.3	2.1 ± 1.9
PDE	5.0 ± 0.1	0.0 ± 0.0
PLC	3.0 ± 0.3	0.0 ± 0.1
PKC	2.8 ± 0.7	0.1 ± 0.4
<i>Pax</i>	2.5 ± 0.6	0.2 ± 0.6

Suga et al., 1999a; 1999b; Koyanagi et al., 1998a; 1998b; Hoshiyama et al., 1998に基づいた。

PTP遺伝子族のサブタイプの多様化がみられた時期には、分子進化速度も速かった。同様の結果は、Gα 遺伝子族・PTK遺伝子族・PDE遺伝子族・PLC及びPKC遺伝子族・Pax遺伝子族でもみられる（表8）。この分子進化速度の上昇については、次節(1-2)で詳しく考察する。

表8 分子進化速度の比較

遺伝子族	V _I	V _{II}	V _I /V _{II}
Gα	1.5 x 10 ⁻⁹	0.24 x 10 ⁻⁹	6.3
PTK	1.4	0.34	4.1
PDE	2.1	0.36	5.8
PLC	0.91	0.26	3.5
PKC	0.43	0.28	1.5
Pax	0.33	0.10	3.3

V_IはI期の分子進化速度を、V_{II}はII期の分子進化速度を表す。Suga et al., 1999a; 1999b; Koyanagi et al., 1998a; 1998b; Hoshiyama et al., 1998に基づいた。

PTP遺伝子族のサブタイプの多様化がみられた時期に起きた生物の進化

PTP遺伝子族のサブタイプの多様化は、菌類と動物の分岐後、海綿動物と真正後生動物の分岐前の時期に集中していることが示された。この時期は、動物進化の初期に相当し、動物が多細胞性を獲得した時期である。したがって、この時期にみられた遺伝子族のサブタイプの多様化と動物の多細胞性の獲得との間には、何らかの関連性のある可能性がある。また、細胞のシグナル伝達に関わる他の遺伝子族でも、同様にサブタイプの多様化が見られた。これらの遺伝子族の多様化が、細胞間コミュニケーションの複雑化に寄与し、多細胞性の獲得につながったのかもしれない。

サブタイプの多様化と多細胞性の進化の関連性を検証する上で、今後、カワカイメンのPTP遺伝子の各サブタイプが、カワカイメンにおいてどのような機能を果たしているのかを解明することは重要である。例えば本論文において、細胞接着に重要な役割を果たす複数の受容体型PTP遺伝子がカワカイメンから単離された。しかし海綿動物は、細胞同士の接着が弱く細胞間の協調もほとんどみられない原始的な多細胞動物である(Margulis & Schwartz, 1998)。カワカイメンの受容体型PTP遺伝子が細胞接着と密接に関わっているのか、すなわち多細胞性に寄与する役割を担っているのか、を明らかにすることは、多細胞性進化の分子基盤の理解につながると思われる。

また、サブタイプの多様化と多細胞性の獲得の関連を検証する上で、立襟鞭毛虫は重要な生物である。この立襟鞭毛虫は、16SrRNAの分子系統樹から多細胞動物に最も近縁な単細胞の原生生物とされている(Wainright et al., 1993)。現在、西頼らによりこの立襟鞭毛虫からPTP遺伝子の探査が試みられている。現時点で、立襟鞭毛虫から6種類のPTP遺伝子が単離され、予備的な解析からPTP遺伝子族のサブタイプのうちいくつかは、立襟鞭毛虫と多細胞動物の分岐前に既に作り出されていたことが示唆されている。このことは、遺伝子重複が起きて重複遺伝子が多細胞性獲得に関わる機能確立までの過渡期に、立襟鞭毛虫が分岐したことを意味するのかもしれない。あるいは、サブタイプの多様化と多細胞性の進化との間には関連性はないのかもしれない。今後、立襟鞭毛虫のPTP遺伝子のさらなる単離と、系統的位置の確認を含む詳細な分子進化学的解析が重要である。

PTP遺伝子族のサブタイプの多様化がみられない時期に起きた生物の進化

海綿動物の分岐後、刺胞動物などの二胚葉動物と、節足動物や脊索動物などの主要な動物門のほとんどすなわち三胚葉動物が分岐し、多様な動物が出現した。化石記録から、多細胞動物の全ての動物門は先カンブリア時代とカンブリア紀の境（5.4億年前）に急激に進化した、とする説がある(Conway Morris, 1993)。この急速な多細胞動物の進化はカンブリア爆発とよばれる。ただし、全ての動物門が短期間に進化したのかについてはいまだ議論がある。最近の分子時計を利用した分岐年代の推定によると、短期間に多様化した動物は三胚葉動物の動物門を指し、その時期は約7～6億年前のようである(Doolittle et al., 1996; Wray et al., 1996; Nikoh et al., 1997; Ayala et al., 1998; Bromham et al., 1998)。いずれにせよカンブリア紀までには、形態上多様な構造をもった動物が進化した。

この多様な動物が進化した時代には、PTP遺伝子族の新しいサブタイプは作り出されていなかった。サブタイプの創生は約9.4億年前に既に完了している。形態形成により深く関与する分子である*Pax*遺伝子族についても、同様のことが示されている(Hoshiyama et al., 1998)。したがって、カンブリア爆発のような生物進化と遺伝子族のサブタイプの多様化とは、その時期に隔たりのあることから考えて、関連の無いことが示唆される。すなわち、多様な多細胞動物の進化は、新しい遺伝子を作り出すことなく達成されたと考えられる。

その他のカワカイメンのPTP遺伝子

カワカイメンからは、結果に示した10種類のPTP遺伝子以外にも、sPTPR3に近縁なPTP遺伝子が17種類単離された（付録5-2にアライメントを示した）。これらの遺伝子はお互いに、アミノ酸配列で28%～95%の高い一致度を示し、最も近縁な配列同士の間では同義置換数は0.04であった。これらの遺伝子は、カワカイメンの系統で独自に遺伝子重複により作り出されたものと考えられる。他の動物において、一つのサブタイプに属する遺伝子がこのように多く作り出された例は知られていない。カワカイメンにおいてこれらPTP遺伝子がどのような機能を担っているのか、また18種類の遺伝子をどう使い分けているのかは興味深い。

PTP遺伝子の分類

本論文における分子系統解析から、PTP遺伝子族が18種類に分類出来ることが示された。これまでPTPの分類は主にドメイン構造からなされていた(Charbonneau & Tonks, 1992; Brady-Kainay & Tonks, 1995)。そのため、PTPドメイン以外にドメイン構造の存在しないPTP遺伝子は、詳しく分類されていなかった。今回の結果から、PTPN3やPTPR7などのドメイン構造をもたないサブタイプを、新たに分類することができた。ドメイン構造が同じPTPR2AとPTPR9などもその起源が異なり、異なるサブタイプだと考えられた。ドメイン構造と進化的な起源を考慮したこの分類は、機能分類に対応することが期待される。

(表9)

また、PTPR8サブタイプに属する遺伝子のN末端の領域に、未同定のドメインを発見した。この領域は、マウスのRESP-18と呼ばれる主要糖質コルチコイド応答タンパク(Bloomquist et al., 1994)と相同性があり、何らかの機能ドメインであることが示唆される。

(図10)

表9 PTP遺伝子族のサブタイプの分類

サブタイプ	ドメイン構造	番号	遺伝子名	生物	アクセッションナンバー (GenBank)
PTPR1/6	FN-(PTP)2	45	LCA	<i>H. sapiens</i>	Y00638
		46	CD45	<i>Heterodontus francisci</i>	U34750
PTPR2A	(Ig)3-(FN)9-(PTP)2	56	Lar	<i>D. melanogaster</i>	M27700
	(Ig)3-(FN)3-(PTP)2	57	HmLAR1	<i>Hirudo medicinalis</i>	AF017084
	(Ig)3-(FN)8-(PTP)2	58	PTPδ	<i>H. sapiens</i>	L38929
	(Ig)3-(FN)8-(PTP)2	59	LAR	<i>H. sapiens</i>	Y00815
	(Ig)3-(FN)4-(PTP)2	60	CRYPα1	<i>Gallus gallus</i>	L32780
	(Ig)3-(FN)8-(PTP)2	61	PTPσ	<i>H. sapiens</i>	U35234
	(Ig)3-(FN)9-(PTP)2	62	HmLAR2	<i>Hirudo medicinalis</i>	AF017083
PTPR2B	MAM-Ig-(FN)4-(PTP)2	50	ptprt	<i>Xenopus laevis</i>	AF173857
		51	RPTP-p	<i>H. sapiens</i>	AF043644
		52	hR-PTPμ	<i>H. sapiens</i>	X58288
		53	h-PTPκ	<i>H. sapiens</i>	Z70660
		54	hPTP-J	<i>H. sapiens</i>	U73727
		*55	sPTPR2B	<i>Ephydatia fluviatilis</i>	AB019126
PTPR3	(FN)9-PTP	*64	sPTPR3	<i>Ephydatia fluviatilis</i>	AB019128
		65	PTP	<i>Gallus gallus</i>	AJ238216
		66	DEP-1	<i>H. sapiens</i>	U10886
	(FN)8-PTP		(HPTPη)		
	(FN)16-PTP	67	HPTPβ	<i>H. sapiens</i>	X54131
	(FN)10-PTP	68	Ptp10D	<i>D. melanogaster</i>	M80465

表9 続き

	(FN)10-PTP	69	Ptp4E	<i>D. melanogaster</i>	L20894
	(FN)5-PTP	70	PTP-U2 (GLEPP1)	<i>H. sapiens</i>	Z48541
	(FN)5-PTP	71	CRYP-2	<i>Gallus gallus</i>	U65891
	(FN)14-PTP	72	PTPRQ	<i>Rattus norvegicus</i>	AF063249
	(FN)10-(PTP)2	73	Esp	<i>Mus musculus</i>	U36488
	(FN)8-PTP	74	SAP-1	<i>H. sapiens</i>	D15049
PTPR4	(PTP)2	47	HPTPε	<i>H. sapiens</i>	X54134
		48	RPTPα	<i>H. sapiens</i>	M34668
		49	PTPα	<i>Xenopus laevis</i>	U09135
		*63	sPTPR4	<i>Ephydatia fluviatilis</i>	AB019125
PTPR5	CA-FN-(PTP)2	*7	sPTPR5	<i>Ephydatia fluviatilis</i>	AB019127
		40	CPTPζ	<i>Gallus domesticus</i>	L27625
		41	PTPζ (RPTPβ)	<i>H. sapiens</i>	M93426
		42	RPTPγ	<i>H. sapiens</i>	L09247
	(FN)2-(PTP)2	43	PTPγ	<i>Gallus gallus</i>	U38349
		44	Ptp99A	<i>D. melanogaster</i>	M80539
PTPR7	PTP	36	STEP	<i>H. sapiens</i>	U27831
		37	HPCPTP1	<i>H. sapiens</i>	D64053
		38	LC-PTP (HePTP)	<i>H. sapiens</i>	D11327
		*39	sPTPR7	<i>Ephydatia fluviatilis</i>	AB019131
PTPR8	unknown-PTP	29	IA-2	<i>H. sapiens</i>	L18983
		30	IAR	<i>H. sapiens</i>	AF007555
PTPR9	(Ig)2-(FN)3-(PTP)2	6	Ptp69D	<i>D. melanogaster</i>	M27699
PTPN1	CRALBP-PTP	21	PTPX1	<i>Xenopus laevis</i>	L33098
		22	PTPX10	<i>Xenopus laevis</i>	L33099
		23	PTP MEG2	<i>H. sapiens</i>	M83738
		*24	sPTPN1	<i>Ephydatia fluviatilis</i>	AB019129
PTPN2	PTP(-PEST)	25	70zpep	<i>H. sapiens</i>	AF077031
		26	PTP G1 (PTP-PEST)	<i>H. sapiens</i>	D13380
		27	BDP1	<i>H. sapiens</i>	X79568
		*28	sPTPN2	<i>Ephydatia fluviatilis</i>	AB019130
PTPN3	PTP	8	PTPN1 (PTP1B)	<i>H. sapiens</i>	M33689
		9	PTP1B	<i>Gallus gallus</i>	U86410
		10	PTPRF	<i>H. sapiens</i>	M25393
		*11	sPTPN3	<i>Ephydatia fluviatilis</i>	
		12	Ptp61F	<i>D. melanogaster</i>	L11253
PTPN4	PDZ-ERM-(PDZ)5-PTP	13	PTP-BAS (FAP-1, hPTP1E, PTP L1)	<i>H. sapiens</i>	D21209

表9 続き

PTPN5	PTP	14	DKFZp566K0524	<i>H. sapiens</i>	AL050040
PTPN6	(SH2)2-PTP	31	PTP1D	<i>H. sapiens</i>	X70766
			(SH-PTP2, SH-PTP3, PTP2C, SHP-2)		
		32	SH-PTP2	<i>Xenopus laevis</i>	U15287
		33	csw	<i>D. melanogaster</i>	U19909
		*34	sPTPN6	<i>Ephydatia fluviatilis</i>	AB019132
		35	hPTP1C	<i>H. sapiens</i>	U15528
			(SH-PTP1, SHP-1, HCP)		
PTPN7	ERM-PDZ-PTP	15	PTPH1	<i>H. sapiens</i>	M64572
		16	PTP MEG	<i>H. sapiens</i>	M68941
PTPN8	ERM-PTP	17	pez	<i>H. sapiens</i>	X82676
		18	PTPD1	<i>H. sapiens</i>	X79510
		*19	sPTPN8	<i>Ephydatia fluviatilis</i>	AB019133
PTPN9	PTP	20	DKFZp564F0923	<i>H. sapiens</i>	AL110210

番号は図7の系統樹上の番号に対応する。*は遺伝子配列を決定したカワカイメンの遺伝子。括弧内に遺伝子の別名を記す。省略は以下である。*H. sapiens*, *Homo sapiens*; *D. melanogaster*, *Drosophila melanogaster*; PTP, protein tyrosine phosphatase domain; Ig, immunoglobulin domain; FN, fibronectin type III domain; MAM, Meprin A5 PTPu domain; CA, carbonic anhydrase-like domain; CRALBP, cellular retinaldehyde binding protein homology domain; PEST, PEST region; PDZ, PSD-95 DigA ZO-1 (GLGF) domain; ERM, ezrin radixin moesin (band 4.1-like) domain; SH2, src homology 2 domain; unknown, unknown domain (図10参照)。ドメイン構造は Streuli (1996) and Stoker (1996)に基づいた。

1.mouse RESP18 (L34214)
2.human IAR (AF007555)
3.human IA-2 (L18983)

(36) GQGQVGSEQLWTFQGLIASVFQYLQLIFHQIVPEGMFWADDIAYELMTKKVEHLSRLHPQYP (77)
(58) GRCQKVPAMDFYRYEVSPVALQRLRVALQKLSGTGFTWQDDYTQYVMDQELADLPKTYLRRP (859)
(59) GQCQVGVGQARPLLQVTSPVLQRLQGVLRQLMSQGLSWHDDL TQYVISQEMERIPRLRPPEP (894)
* * * * *

図10 PTPR8サブタイプの未同定のドメインのアライメント
*は全ての配列で保存しているアミノ酸を示す。括弧内はその領域に存在するアミノ酸の総数。

1.2 PTP遺伝子族のアイソフォームの多様化

動物はその進化の初期の段階で、18種類の機能の異なるPTP遺伝子族のサブタイプを作り出していた。そして、それ以降新しいサブタイプはほとんど作り出されていなかった。一方、多くのサブタイプはその後脊椎動物の系統において遺伝子が増えている。これらの遺伝子は同じドメイン構造を保持することから、基本的な機能は共通していると考えられる。本論文ではこれをアイソフォームと呼ぶことにする。アイソフォーム間の差異は、表10のようにしばしば発現する組織に見られる。したがって、各サブタイプは複数のアイソフォームを作り出し、それぞれのアイソフォームは異なる組織で発現するように進化した、と考えられる。

これらのアイソフォームは、海綿動物と真正後生動物の分岐後、脊椎動物に至る系統において、遺伝子重複により作り出されたと考えられる。それでは、PTP遺伝子族のアイソフォームが作り出された時期はいつだろうか。このことを明らかにする目的で、脊椎動物に最も近縁な無脊椎動物である頭索動物(Willmer 1990; Wada and Satoh, 1994)のナメクジウオ、脊椎動物の中で最初に分岐した無顎類(Romer & Parsons, 1986)のヌタウナギ、有顎類の中で最初に分岐した軟骨魚類(Romer & Parsons, 1986; Colbert & Morales, 1991)のポタモトリゴンモトロ（淡水エイ）から、PTP遺伝子の探査、及び分子進化学的解析を行った（図11）。

表10 PTP遺伝子の各サブタイプに属するアイソフォームの組織特異的発現

サブタイプ名	遺伝子名	発現する組織	参考文献
PTPR2A	LAR	広く発現	a
	PTPδ	広く発現	a
	PTPσ	脳	a
PTPRR2B	PTPμ	肺	b
	PTPκ	肝臓・腎臓	b
	PTPλ	脳、肺、腎臓	b
	PTPρ	中枢神経	b
PTPR5	PTPγ	広く発現	c
	PTPζ	脳	d
PTPN6	SHP-1	造血細胞	e
	SHP-2	広く発現	e

参考文献をもとに、哺乳類の遺伝子で組織特異的に発現するPTP遺伝子をまとめた。参考文献の略号は以下を示す。a, Pulido et al., 1995; b, 水野、1998; c, Barnea et al., 1993; d, Krueger & Saito, 1992; e, Streuli, 1996.

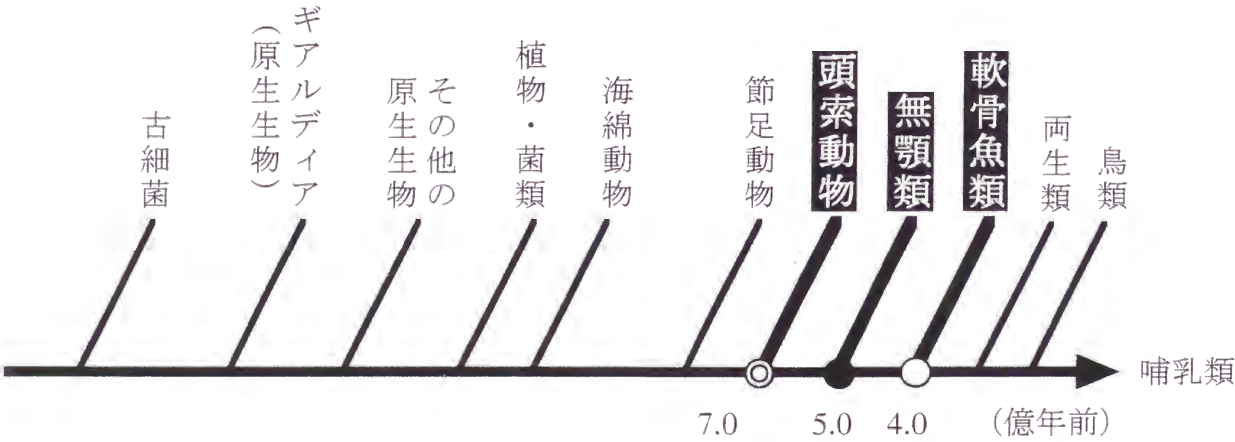


図11 頭索動物、無顎類、軟骨魚類の系統的位置

本節で注目する頭索動物、無顎類、軟骨魚類の系統的位置を模式的に示した。枝の長さは概ね時間を反映しているが正確ではない。数字は分岐年代を示す(Dickerson, 1971; Dayhoff, 1978)。節足動物及び頭索動物の分岐時期は接近しており約7億年前とされる(Nikoh et al., 1997)。

結果

ナメクジウオ・ヌタウナギ・エイのPTP遺伝子のクローニング

RT-PCRによりナメクジウオ・ヌタウナギ・エイからPTP遺伝子の探査を行った。その結果、ナメクジウオから9種類、ヌタウナギから11種類、エイから13種類の、計33種類のPTP遺伝子の単離に成功した（塩基配列及び推定アミノ酸配列は付録2-2に示した）。塩基配列を決定した領域は、PTPドメインのN末側約1/3から終止コドンまでに相当する。（エイのryPTPN6b及びryPTPN6cについては、全PTPドメインの塩基配列が決定された。）

ナメクジウオのPTP遺伝子は他の動物のPTP遺伝子と30～94%、ヌタウナギのPTP遺伝子は27～95%、エイのPTP遺伝子は29～97%、の一致度を示した。また、単離した全てのPTP遺伝子に、PTP signature motif 'HCXXGXXRS/T' (Denu et al., 1996)が保存していた（付録2-2参照）。したがって、これらの遺伝子はPTPとして機能しているものと思われる。

単離した遺伝子のサブタイプの同定

単離したナメクジウオ・ヌタウナギ・エイのPTP遺伝子が、どのサブタイプに属するかを調べる目的で、既存のデータベース中に存在するPTP遺伝子と、単離されたナメクジウオ・ヌタウナギ・エイのPTP遺伝子を併せて、PTP遺伝子族全体の分子系統樹を推定した（図12）。

その結果、単離した33遺伝子のうち32遺伝子は、前節1-1で定義したいずれかのサブタイプに分類された。単離したPTP遺伝子（白抜き文字で示した）と各サブタイプに属する遺伝子は、高いブートストラップ確率で単系統群を形成しており、この分類が確実であることを示している。また図12の分子系統樹において、ナメクジウオのamPTPR10は、どのサブタイプにも属さなかった。amPTPR10はPTPドメインを二つ有しており、新しい受容体型PTP遺伝子のサブタイプであることが示唆される。ただし、受容体型のPTP遺伝子のみを選び出し、系統樹推定に用いるアライメント領域を広げて推定した分子系統樹からは、PTPR2Bサブタイプに属する可能性も示されている。

PTPR4、PTPN3、PTPR5、PTPN6、PTPR2Aの5つのサブタイプにおいては、ナメクジウオ・ヌタウナギ・エイの3種全てからPTP遺伝子が単離された。そこで、この5サブタイプについて、遺伝子重複によりアイソフォームが作り出された時期の詳細な解析を行った。

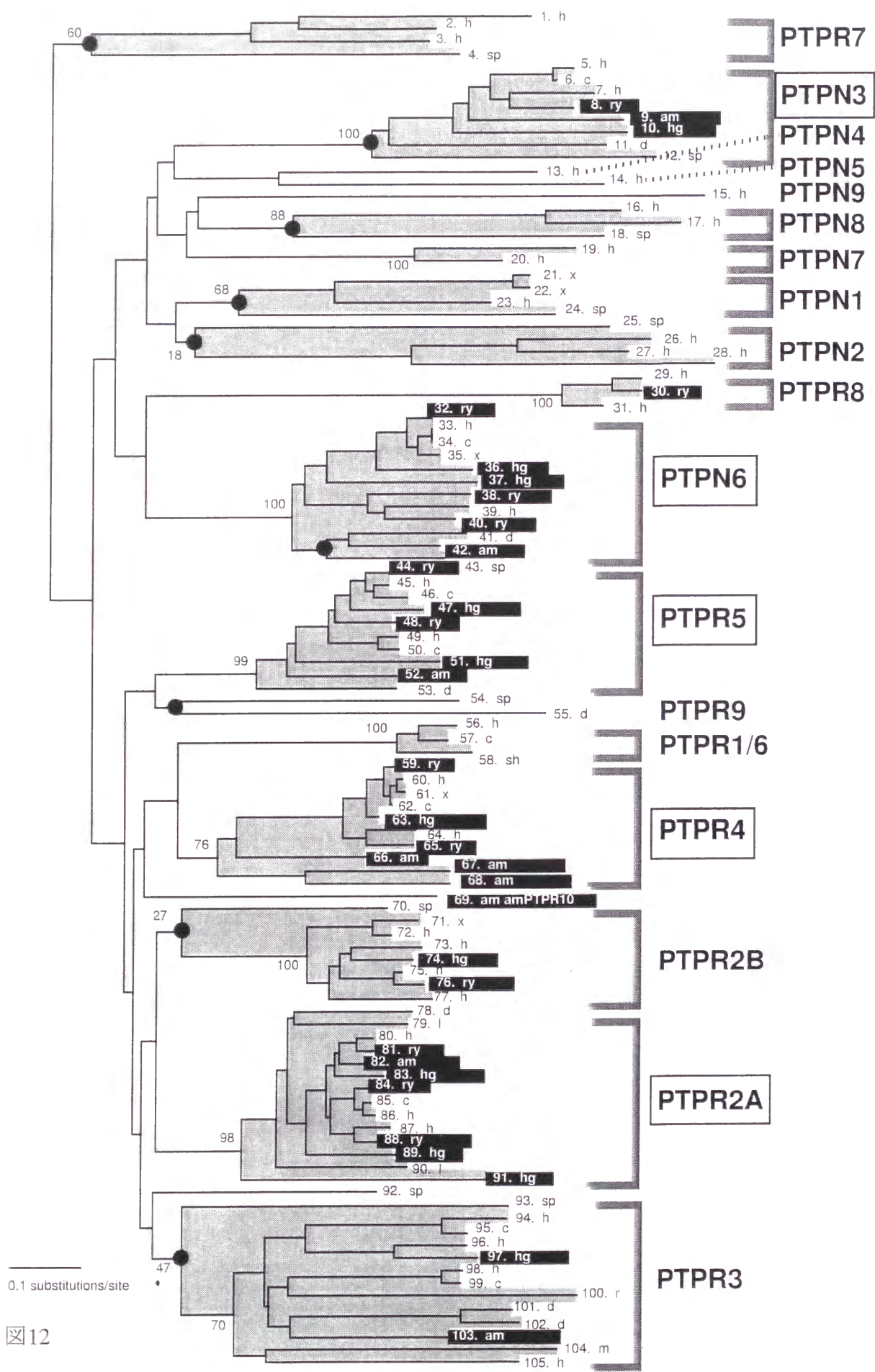


図12 PTP遺伝子族の分子系統樹

PTPドメインの約3分の2に相当するアライメント領域（136アミノ酸座位）に基づき、近隣結合法により分子系統樹を推定した。アライメントは付録5-1に準ずる。遺伝子配列を決定したナメクジウオ・ヌタウナギ・エイの遺伝子は白抜き文字で示した。各系統の記号は以下を表す。h, human; m, mouse; r, rat; c, chicken; x, *Xenopus*; sh, shark; ry, ray; hg, hagfish; am, amphioxus; d, *Drosophila*; l, leech; s, sponge. 系統樹推定に用いた遺伝子の名前及びGenBankのアクセッションナンバーは付録3-1に示した。同じサブタイプに属する遺伝子群は網掛けで示し、サブタイプ名を右側に記した。アイソフォームの作り出された時期について詳細な解析を行ったサブタイプは四角で囲った。ノード付近に示した数字はブートストラップ確率を示す。

この系統樹は前節1-1の分子系統樹からアウトグループの菌類の遺伝子を除いた無根系統樹である。無根系統樹とした理由は、遠縁の菌類の遺伝子配列を除くことで、アライメント領域を広げ、より信頼性の高い系統樹が得られると考えたからである。（アウトグループとして菌類の遺伝子を含めた場合にも、遺伝子配列を決定した遺伝子のサブタイプの分類結果は変わらなかった。）また、海綿動物と真正後生動物の分岐を黒丸で示した。この系統樹からも、サブタイプの多様化が海綿動物と真正後生動物の分岐前に起きているという前節1-1の結果が示されている。

PTP遺伝子族のアイソフォームが作り出された時期の推定

A. PTPR4サブタイプ

脊椎動物には、PTPR4サブタイプにPTP α とPTP ϵ の2種類のアイソフォームが存在する。本論文では新たにナメクジウオから3種類、ヌタウナギから1種類、エイから2種類の、PTPR4サブタイプに属する遺伝子を単離した。これらの遺伝子を併せて、発見的探索による最尤法を用いて、PTPR4サブタイプの分子系統樹を推定した（図13）。アウトグループには、PTPR1/6サブタイプに属する遺伝子を用いた（網掛け）。

エイのryPTPR4b及びヌタウナギのhgPTPR4は、ヒト・ニワトリ・アフリカツメガエルのPTP α 遺伝子と単系統群を形成しており、その分岐の順序は生物種の分岐を反映している。したがって、エイのryPTPR4b及びヌタウナギのhgPTPR4は、ヒト・ニワトリ・アフリカツメガエルのPTP α とオーソログスな関係にある遺伝子だと考えられる。同様に、エイのryPTPR4aはヒトのHPTP ϵ とクラスターを成し、PTP ϵ アイソフォームに分類される。一方、ナメクジウオの3遺伝子は分子系統樹上で単系統群を形成している。このことはこの3遺伝子が、ナメクジウオの系統で独自に遺伝子重複により作り出されたことを示している。このナメクジウオの3遺伝子は、PTP α アイソフォーム及びPTP ϵ アイソフォームとオーソログスな関係にある遺伝子である。

この分子系統樹から、PTP α とPTP ϵ の2種類のアイソフォームが作り出された時期が推定できる。PTP α アイソフォームとPTP ϵ アイソフォームを作り出した遺伝子重複（菱形で示す）は、ナメクジウオと脊椎動物の分岐（二重丸）後、ヌタウナギと有顎類の分岐（黒丸）前に起きている。

発見的探索による最尤法で得られた最尤系統樹は、解を網羅的に探索をしていないことから、局所解である危険性がある。すなわち、真の最尤系統樹を見落としている可能性がある。そこでこの結果を確かめるため、局所ブートストラップ確率 (LBP) が95%以上となる分岐を正しいと仮定し、網羅的探索による最尤法で系統樹推定を行った。その結果、図13と同じ樹形の分子系統樹が最尤系統樹となり、他の樹形は統計的に棄却された。したがって、この系統樹が真の最尤系統樹だと考えられる。

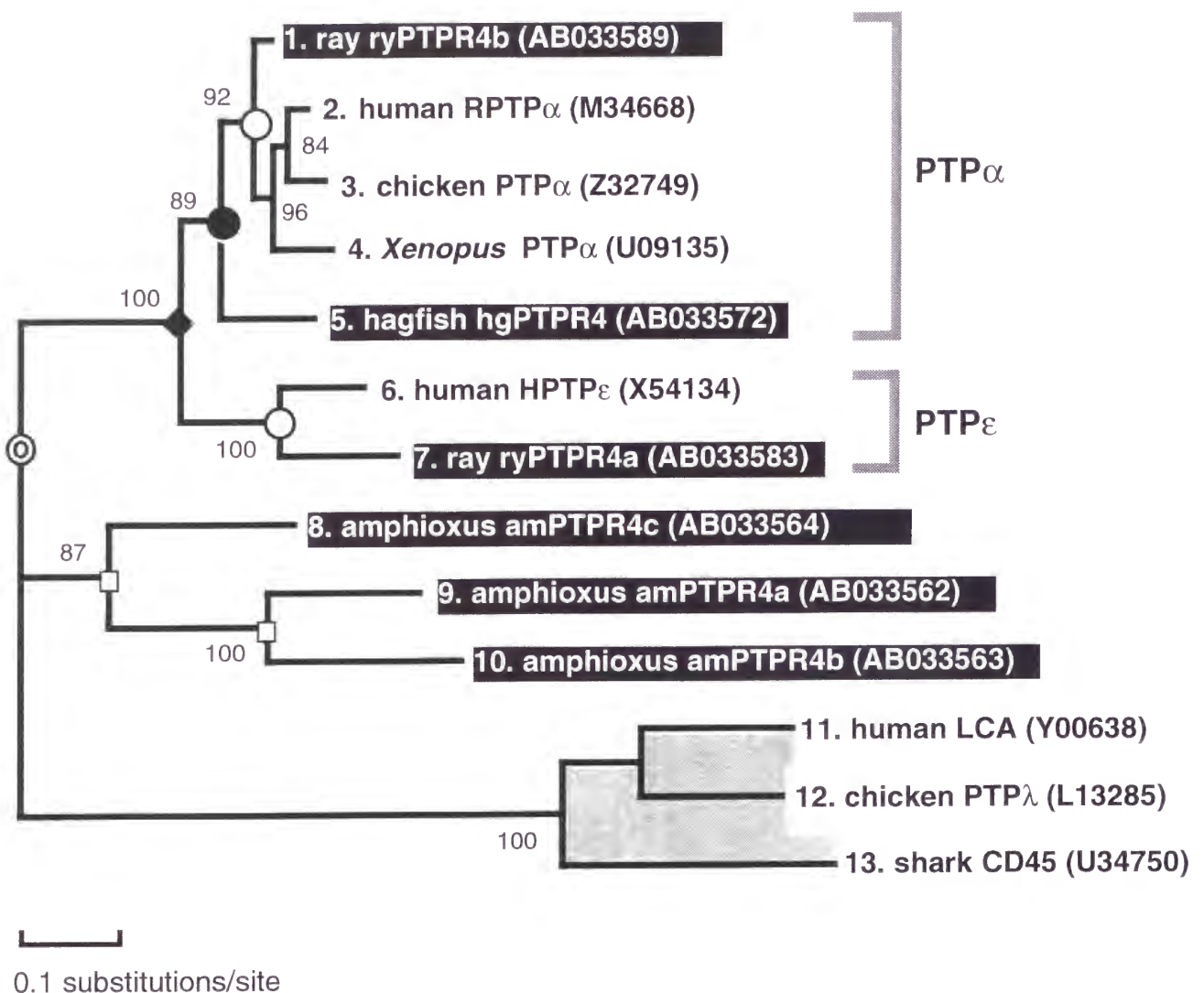


図13 PTPR4サブタイプの分子系統樹

N末側のPTPドメイン約3分の2とC末側のPTPドメイン全てを含むアライメント領域（416アミノ酸座位）に基づき、発見的探索による最尤法で最尤系統樹を推定した。アライメントは付録5-3に示した。アウトグループにはPTPR1/6サブタイプに属する遺伝子群を用いた（網掛け）。本論文で決定した配列は白抜き文字で示した。配列には生物種名、遺伝子名、GenBankのアクセッションナンバーを記載し、アイソフォームの名前を右側に記した。ノード付近に示した数字は局所ブートストラップ確率(LBP)である。LBPが95%以上の分岐を正しいと仮定し、網羅的探索による最尤法から得られた最尤系統樹もこの樹形と一致した。また、他の樹形の尤度は、この最尤系統樹の尤度と1標準誤差以上の差があり棄却された。

二重丸は頭索動物と脊椎動物の分岐を、黒丸は無顎類と有顎類の分岐を、白丸は軟骨魚類とその他の有顎類の分岐を表す。また、菱形は脊椎動物の系統で起きた遺伝子重複、四角は無脊椎動物の系統で起きた遺伝子重複である。

B. PTPN3サブタイプ

脊椎動物には、PTPN3サブタイプにPTPN1とPTPRFの2種類のアイソフォームが存在する。本論文では新たにナメクジウオから1種類、ヌタウナギから1種類、エイから1種類の、PTPN3サブタイプに属する遺伝子を単離した。これらの遺伝子を併せて、網羅的探索による最尤法を用いて、PTPN3サブタイプの分子系統樹を推定した。アウトグループには、PTPN3サブタイプに属するショウジョウバエのPtp61F遺伝子を用いた。その結果、図14の分子系統樹が最尤系統樹となり、他の樹形は統計的に棄却された。

図14の分子系統樹から、エイのryPTPN3はPTPRFアイソフォームに分類された。ヌタウナギのhgPTPN3及びナメクジウオのamPTPN3は、PTPN1アイソフォーム及びPTPRFアイソフォームとオーソログな関係にある遺伝子である。

この2種類のアイソフォームを作り出した遺伝子重複（菱形で示す）は、PTPR4サブタイプと同様、ナメクジウオと脊椎動物の分岐（二重丸）後に起きている。しかし、PTPR4サブタイプと異なり、ヌタウナギと有顎類の分岐（黒丸）後でもあった。また少なくとも、エイとその他の有顎類（白丸）との分岐前である。

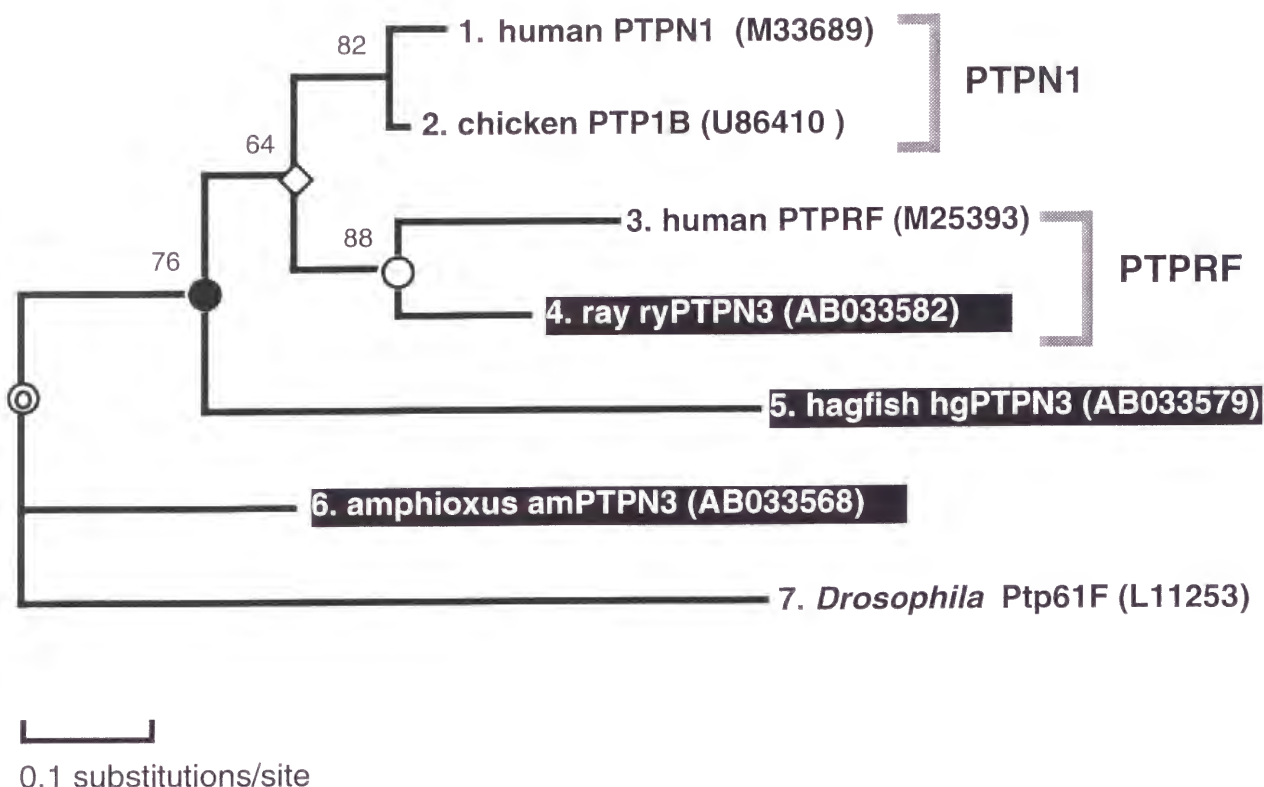


図14 PTPN3サブタイプの分子系統樹

PTPドメインの約3分の2に相当するアライメント領域（180アミノ酸座位）に基づき、網羅的探索による最尤法で最尤系統樹を推定した。アライメントは付録5-4に示した。アウトグループにはPTPN3サブタイプに属するショウジョウバエのPtp61F遺伝子を用いた。本論文で決定した配列は白抜き文字で示した。配列には生物種名、遺伝子名、GenBankのアクセッションナンバーを記載し、アイソフォームの名前を右側に記した。ノード付近に示した数字はブートストラップ確率である。また、他の樹形の尤度は、この最尤系統樹の尤度と1標準誤差以上の差があり棄却された。

二重丸は頭索動物と脊椎動物の分岐を、黒丸は無顎類と有顎類の分岐を、白丸は軟骨魚類とその他の有顎類の分岐を、菱形は遺伝子重複を表す。

C. PTPR5サブタイプ

脊椎動物には、PTPR5サブタイプにPTP γ とPTP ζ の2種類のアイソフォームが存在する。本論文では新たにナメクジウオから1種類、ヌタウナギから2種類、エイから2種類の、PTPR5サブタイプに属する遺伝子を単離した。これらの遺伝子を併せて、網羅的探索による最尤法を用いて、PTPR5サブタイプの分子系統樹を推定した。アウトグループには、PTPR5サブタイプに属するショウジョウバエのPtp99A遺伝子を用いた。その結果図15の系統樹が最尤系統樹となった。

図15の分子系統樹から、エイのryPTPR5b及びヌタウナギのhgPTPR5aはPTP γ アイソフォームに分類された。また、エイのryPTPR5aはPTP ζ アイソフォームに分類された。ナメクジウオのamPTPR5は、PTP γ アイソフォーム及びPTP ζ アイソフォームとオーソログスな関係にある遺伝子である。

この2種類のアイソフォームを作り出した遺伝子重複（菱形で示す）は、PTPR4サブタイプと同様、ナメクジウオと脊椎動物の分岐（二重丸）後、ヌタウナギと有顎類の分岐（黒丸）前に起きている。

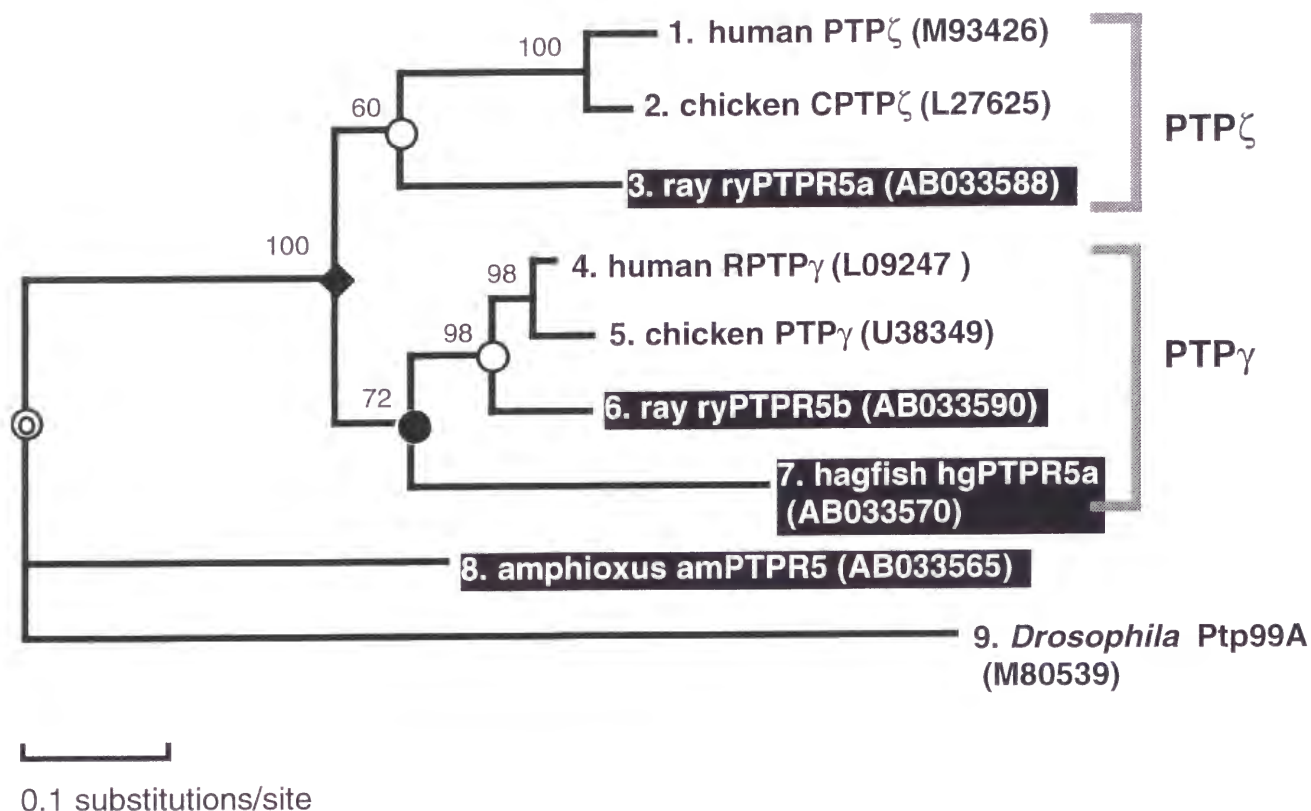


図15 PTPR5サブタイプの分子系統樹

N末側のPTPドメイン約3分の2とC末側のPTPドメイン全てを含むアライメント領域（435アミノ酸座位）に基づき、網羅的探索による最尤法で最尤系統樹を推定した。アライメントは付録5-5に示した。アウトグループにはPTPR5サブタイプに属するショウジョウバエのPtp99A遺伝子を用いた。本論文で決定した配列は白抜き文字で示した。配列には生物種名、遺伝子名、GenBankのアクセッションナンバーを記載し、アイソフォームの名前を右側に記した。ノード付近に示した数字はブートストラップ確率である。この最尤系統樹と尤度の差が1標準誤差範囲内である樹形が他に3通り（135135通り中）存在した。

ヌタウナギからはPTPR5サブタイプに属する遺伝子が2種類単離されたが、このうちhgPTPR5bは他の遺伝子と異なりPTPドメインを一つしか持たなかった。この遺伝子を含めると系統樹推定に用いるアライメント領域が半減し、系統樹推定の信頼性が低下する。したがって、結論が変わらないことを確認した上で、この遺伝子は除外して分子系統樹を推定した。

二重丸は頭索動物と脊椎動物の分岐を、黒丸は無顎類と有顎類の分岐を、白丸は軟骨魚類とその他の有顎類の分岐を、菱形は遺伝子重複を表す。

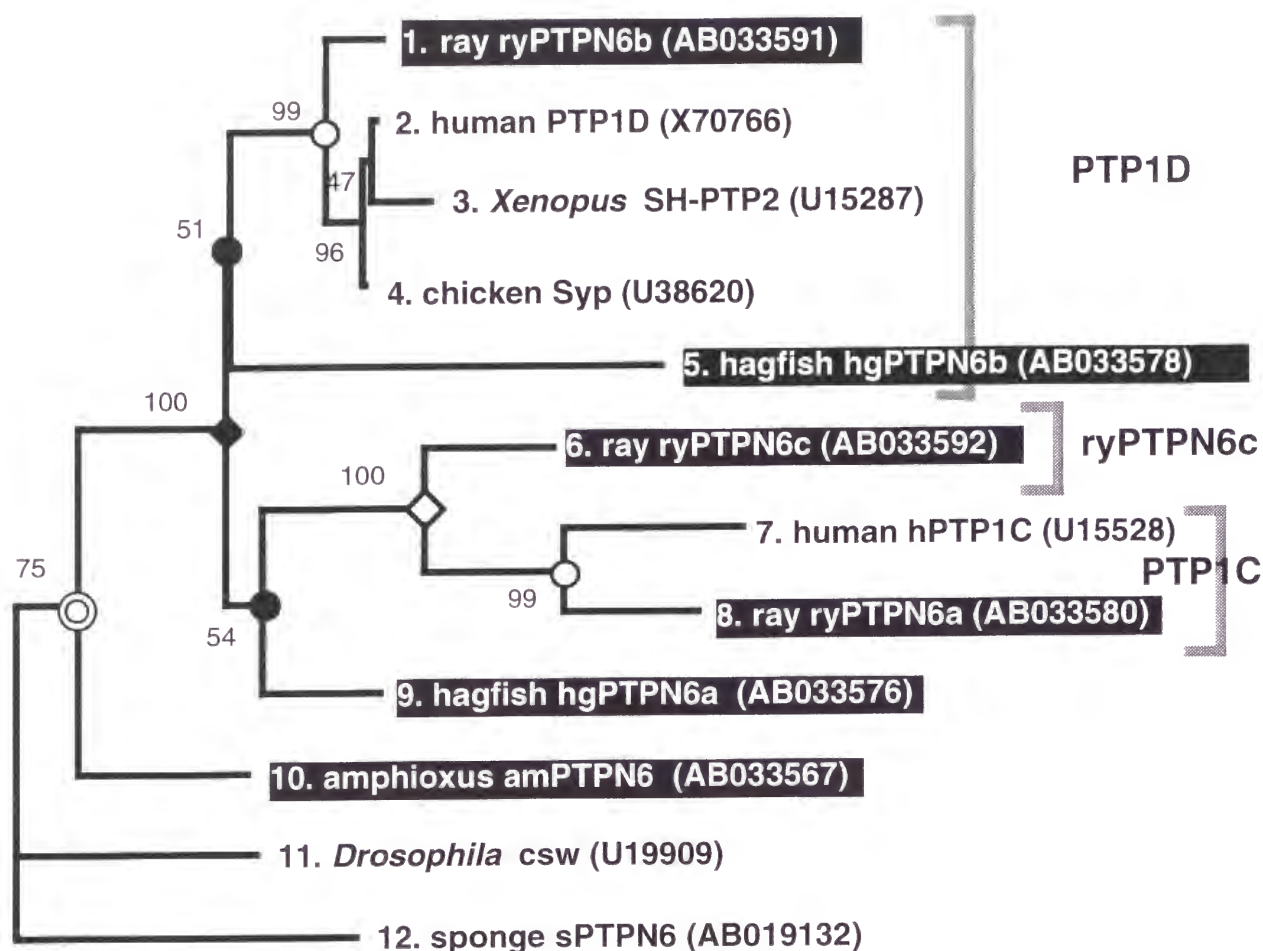
D. PTPN6サブタイプ

脊椎動物には、PTPN6サブタイプにPTP1CとPTP1Dの2種類のアイソフォームが存在する。本論文では新たにナメクジウオから1種類、ヌタウナギから2種類、エイから3種類の、PTPN6サブタイプに属する遺伝子を単離した。これらの遺伝子を併せて、発見的探索による最尤法を用いて、PTPN6サブタイプの分子系統樹を推定した（図16）。アウトグループには、PTPN6サブタイプに属するカワカイメン及びショウジョウバエの遺伝子を用いた。

図16の分子系統樹から、エイのryPTPN6aはPTP1Cアイソフォームに分類された。また、エイのryPTPN6b及びヌタウナギのhgPTPN6bはPTP1Dアイソフォームに分類された。エイのryPTPN6cは、新しいアイソフォームだと考えられる。ヌタウナギのhgPTPN6aは、PTP1Cアイソフォーム及びryPTPN6cアイソフォームと、ナメクジウオのamPTPN6は3種のアイソフォームとそれぞれオーソログな関係にある遺伝子である。

この3種類のアイソフォームを作り出した遺伝子重複2回（菱形で示す）のうち、1回はナメクジウオと脊椎動物の分岐（二重丸）後、ヌタウナギと有顎類の分岐（黒丸）前に起きている。もう1回は、ヌタウナギと有顎類の分岐後、エイと他の有顎類の分岐（白丸）前に起きている。

図16の系統樹は発見的探索による最尤法で得られた系統樹であるため、LBPが99%以上である分岐を仮定して、網羅的探索による最尤法でも系統樹推定を行った。その結果、図16と同じ樹形の分子系統樹が最尤系統樹となった。



0.1 substitutions/site

図16 PTPN6サブタイプの分子系統樹

PTPドメイン約3分の2に相当するアライメント領域（182アミノ酸座位）に基づき、発見的探索による最尤法で推定された最尤系統樹を示した。アライメントは付録5-6に示した。アウトグループにはPTPN6サブタイプに属するカワカイメン及びショウジョウバエの遺伝子を用いた。本論文で決定した配列は白抜き文字で示した。配列には生物種名、遺伝子名、GenBankのアクセッションナンバーを記載し、アイソフォームの名前を右側に記した。ノード付近に示した数字はLBPである。LBPが99%以上の分岐を正しいと仮定し、網羅的探索による最尤法から得られた最尤系統樹もこの樹形と一致した。ただし、この最尤系統樹と尤度の差が1標準誤差範囲内である樹形が他に9通り（105通り中）存在した。

二重丸は頭索動物と脊椎動物の分岐を、黒丸は無顎類と有顎類の分岐を、白丸は軟骨魚類とその他の有顎類の分岐を示す。菱形は遺伝子重複を表し、黒の菱形は、頭索動物と脊椎動物の分岐後、無顎類と有顎類の分岐前に起きた遺伝子重複を、白の菱形は、無顎類と有顎類の分岐後、軟骨魚類と他の有顎類との分岐前に起きた遺伝子重複を示す。

E. PTPR2Aサブタイプ

脊椎動物には、PTPR2AサブタイプにLAR、PTP δ とPTP σ の3種類のアイソフォームが存在する。本論文では新たにナメクジウオから1種類、ヌタウナギから3種類、エイから3種類の、PTPR2Aサブタイプに属する遺伝子を単離した。これらの遺伝子を併せて、発見的探査による最尤法を用いて、PTPR2Aサブタイプの分子系統樹を推定した（図17）。アウトグループには、PTPR2Aサブタイプに属する節足動物及び環形動物の遺伝子を用いた。

図17の分子系統樹から、エイのryPTP2AaはLARアイソフォームに、エイのryPTP2AcはPTP σ アイソフォームに、それぞれ分類された。ryPTP2Abは新しいアイソフォームと考えられる。しかし、後述する網羅的探査による最尤系統樹推定からは、ヒトのPTP δ と単系統群を形成する樹形も得られており、PTP δ に分類される可能性も残る。ナメクジウオのamPTPR2A及びヌタウナギのhgPTPR2AaとhgPTPR2Abは、4種のアイソフォームとオーソログスな関係にある遺伝子である。

この4種類のアイソフォームを作り出した遺伝子重複3回（菱形で示す）は全て、ヌタウナギと有顎類の分岐（黒丸）後、エイと他の有顎類の分岐（白丸）前に起きている。

図17の系統樹は発見的探査による最尤法で得られた系統樹であるため、LBPが99%以上である分岐を仮定して、網羅的探査による最尤法でも系統樹推定を行った。その結果、図17と同じ樹形の分子系統樹が最尤系統樹となった。

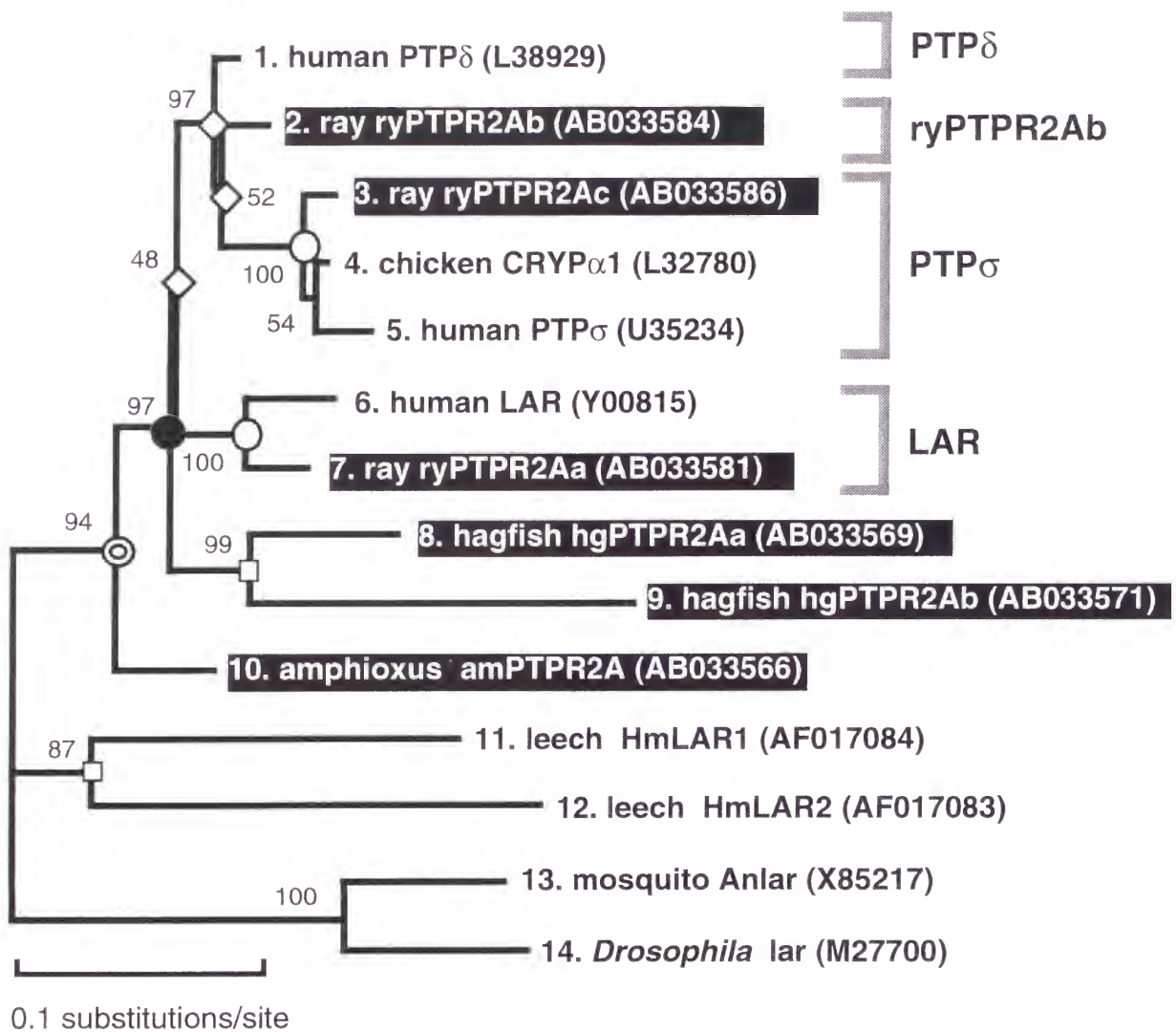


図17 PTPR2Aサブタイプの分子系統樹

N末側のPTPドメイン約3分の2とC末側のPTPドメイン全てを含むアライメント領域（458アミノ酸座）に基づき、発見的探索による最尤法で推定された最尤系統樹を示した。アライメントは付録5-7に示した。アウトグループにはPTPR2Aサブタイプに属する節足動物及び環形動物の遺伝子を用いた。本論文で決定した配列は白抜き文字で示した。配列には生物種名、遺伝子名、GenBankのアクセッションナンバーを記載し、アイソフォームの名前を右側に記した。ノード付近に示した数字はLBPである。LBPが99%以上の分岐を正しいと仮定し、網羅的探索による最尤法から得られた最尤系統樹もこの樹形と一致した。ただし、最尤系統樹と尤度の差が1標準誤差範囲内である樹形が他に8通り（945通り中）存在した。

ヌタウナギからはPTPR2Aサブタイプに属する遺伝子が3種類単離されたが、このうちhgPTPR4cは進化速度が他の遺伝子と比べ非常に高く、この遺伝子を含めると系統樹推定の信頼性が低下すると考え、除外して分子系統樹を推定した。ただし、この遺伝子を含めた網羅的探索による最尤法から、このhgPTPR4cを含むヌタウナギの3遺伝子は単系統群を形成することが、有意（1標準誤差範囲内）に示された。したがってこのヌタウナギの遺伝子を除外しても結論に影響はない。

二重丸は頭索動物と脊椎動物の分岐を、黒丸は無顎類と有顎類の分岐を、白丸は軟骨魚類とその他の有顎類の分岐を示す。菱形は脊椎動物の系統で起きた遺伝子重複、四角は無脊椎動物の系統で起きた遺伝子重複である。

PTP遺伝子族のアイソフォームを作り出す遺伝子重複が起きた時期の再検討

以上5つのサブタイプにおいて、アイソフォームを作り出した遺伝子重複が起きた時期は全て、ナメクジウオと脊椎動物の分岐後、エイとその他の有顎類の分岐前であった。遺伝子重複が起きた時期が、ヌタウナギと有顎類の分岐前なのか後なのかについては、サブタイプによって異なる結果が得られた。

PTPR5、PTPN6、PTPR2Aの3つのサブタイプについては、網羅的探索による最尤法の結果、図15～17で示した最尤系統樹の他に、一標準誤差範囲内に棄却できない複数の樹形が存在した。したがって、一標準誤差範囲内に存在する他の樹形が真の系統樹である可能性がある。その場合には、その真の系統樹について、アイソフォームを作り出した遺伝子重複が起きた時期を検討する必要がある。そこで、一標準誤差範囲内にある樹形全てについて、アイソフォームを作り出した遺伝子重複が起きた時期及びその回数を推定し、各樹形のブートストラップ確率で重みづけをして、遺伝子重複が起きた時期について再検討した（材料と方法の章参照）。その結果、図15～17で得られた結果と同様、5つのサブタイプ全てにおいて、頭索動物と脊椎動物の分岐前（III-a期）、あるいは軟骨魚類とその他の有顎類の分岐後（IV期）には、遺伝子重複がほとんど見られないことが確認された。またこのうち、ヌタウナギと有顎類の分岐前（III-b期）に起きた遺伝子重複は計4.0回、これに対してヌタウナギと有顎類の分岐後（III-c期）に起きた遺伝子重複は計3.5回となった（表11）。このことから、ヌタウナギと有顎類の分岐に関しては、それ以前に起きた遺伝子重複とそれ以降に起きたものが、同程度であることが示された。

表11 アイソフォームを作り出す遺伝子重複が起きた時期

サブタイプ	III-a期	III-b期	III-c期	IV期
PTPR4	0.0	1.0	0.0	0.0
PTPN3	0.0	0.0	1.0	0.0
PTPR5	0.0	1.2	0.2	0.0
PTPN6	0.2	1.3	0.7	0.0
PTPR2A	0.0	0.5	1.6	0.0
total	0.2	4.0	3.5	0.0

III-a期はナメクジウオと脊椎動物の分岐前の時期、III-b期はナメクジウオと脊椎動物の分岐後ヌタウナギと有顎類の分岐前の時期、III-c期はヌタウナギと有顎類の分岐後エイとその他の有顎類の分岐前の時期、IV期はエイとその他の有顎類の分岐後現在までを指す。

分子進化速度の比較

PTPN3、PTPR5、PTPN6、PTPR2Aの4つのサブタイプの分子系統樹（図14～図17）からは、アイソフォームが作り出された時期には分子進化速度も高いことが示唆される。節足動物と脊椎動物の分岐後、軟骨魚類とその他の有顎類の分岐前の時期をⅢ期、軟骨魚類とその他の有顎類の分岐後現在までをⅣ期とし、両期の分子進化速度を比較した。節足動物と脊椎動物の分岐は7億年前、軟骨魚類とその他の有顎類の分岐は4.0億年前とした (Dayhoff, 1978)。その結果、サブタイプごとに分子進化速度は異なるものの、Ⅲ期はⅣ期に比べて分子進化速度が速く、平均で約3.9倍になることがわかった（表12）。すなわち、Ⅲ期には遺伝子重複のみでなくアミノ酸の置換も多く起きていることがわかった。

表12 分子進化速度の比較

サブタイプ名	VⅢ	VⅣ	VⅢ/VⅣ
PTPN3	1.0.x10 ⁻⁹	0.67x10 ⁻⁹	1.5
PTPR5	1.2	0.28	4.3
PTPN6	1.3	0.41	3.1
PTPR2A	0.26	0.039	6.7

1アミノ酸座位当たり、1年当たりに起きたアミノ酸の置換数を推定した。VⅢはⅢ期の分子進化速度を、VⅣはⅣ期の分子進化速度を表す。図14～図17の系統樹において、節足動物の分岐の位置を知るためにはさらに遠縁の遺伝子をアウトグループにする必要がある。そこで各サブタイプについて、最も近縁な別のサブタイプをアウトグループとして最尤法により系統樹を推定し、その系統樹に基づき分子進化速度を推定した。PTPR4サブタイプには節足動物のデータが無いため、分子進化速度の比較を行うことが出来なかった。

考察

結果のまとめ

RT-PCRによりナメクジウオ・ヌタウナギ・エイから計33種類のPTP遺伝子の単離に成功した。

単離した遺伝子を含めて分子進化学的解析を行った結果、PTPR2A、PTPR4、PTPR5、PTPN3、PTPN6の5つサブタイプにおいて、アイソフォームの作り出された時期は、頭索動物と脊椎動物の分岐（約7億年前、Nikoh et al., 1997）後、軟骨魚類と他の有顎類の分岐（約4億年前、Dayhoff, 1978）前の間であった（図18）。このうち、無顎類の分岐（約5億年前、Dickerson, 1971）の前後には同程度の遺伝子重複がみられた。

また、アイソフォームを作り出す遺伝子重複が起きた時期は、そうでない時期と比べて、分子進化速度が高かった。

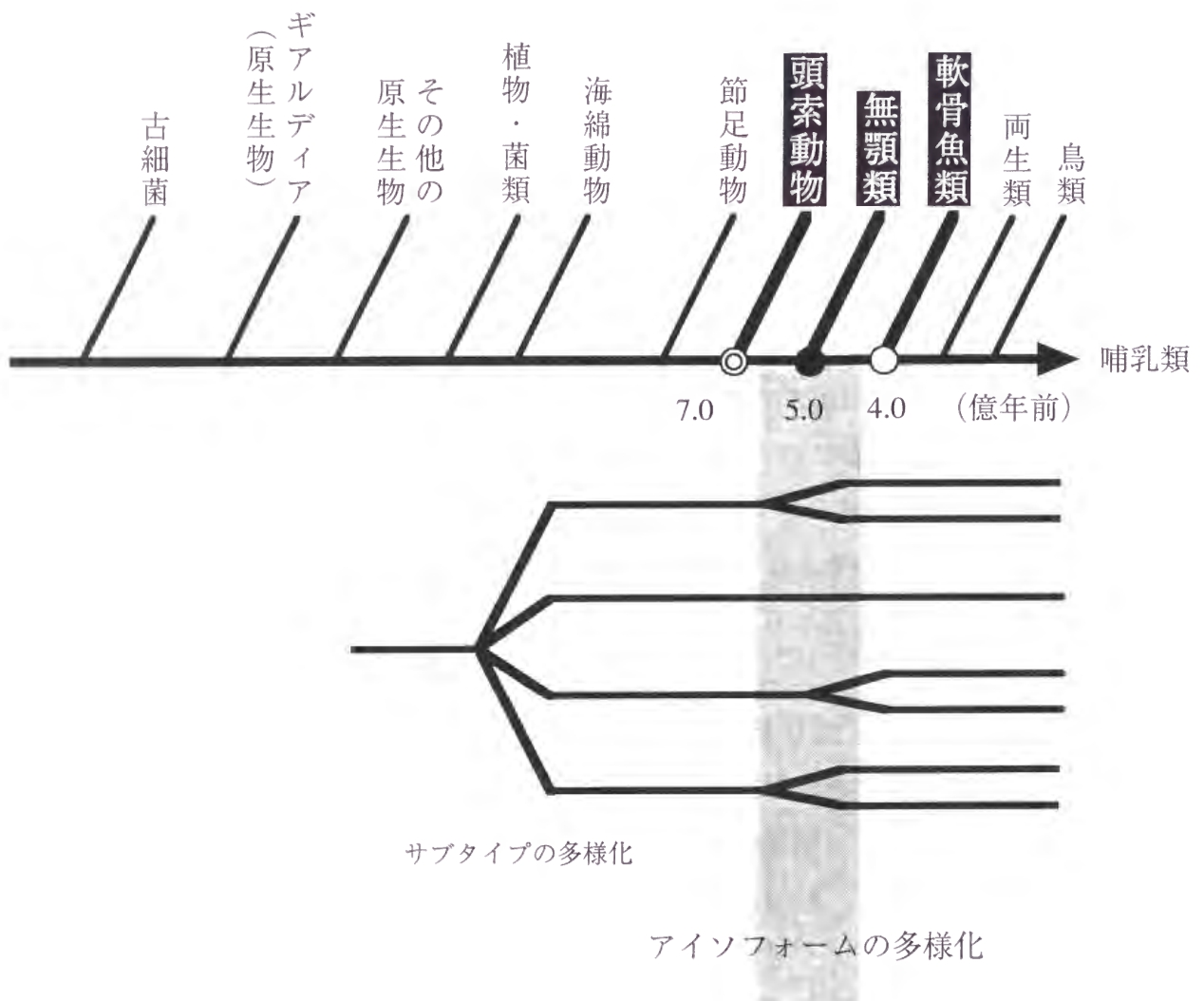


図18 PTP遺伝子族のアイソフォームの多様化のまとめ

PTP遺伝子族のアイソフォームの多様化の時期を模式的に示した。本論文の結果から明らかとなった、アイソフォームの多様化が見られた時期を網掛けで示した。枝の長さは概ね時間を反映しているが正確ではない。数字は分岐年代を示す(Dickerson, 1971; Dayhoff, 1978; Nikoh et al., 1997)。

他の遺伝子族との比較

PTP遺伝子族のアイソフォームが、遺伝子重複により作り出された時期は、無顎類が分岐する時期の周辺だった。同様の結果は、PTK遺伝子族や解糖系の酵素などでもみられる(表13、Suga et al., 1999b)。本論文では新たに、解析した5つのサブタイプ全てについて、アイソフォームを作り出した遺伝子重複が、少なくとも軟骨魚類とその他の有顎類の分岐前に起きたことを示した。

アイソフォームの多様化は、サブタイプの多様化と同様、ある限定した時期に起きていた。一方、様々な遺伝子族でアイソフォームの多様化が見られることは、サブタイプの多様化が細胞のシグナル伝達に関わる遺伝子族が中心であることと、異なる傾向である。この脊椎動物進化の初期には、染色体レベルの大きな重複、あるいはゲノム全体の重複すなわち四倍体化が起きたことが、HOXクラスターの解析(Holland & Garcia-Fernandez, 1996 for review)、脊椎動物の異なる種間でのホモログ遺伝子の染色体上の位置の比較(Postlethwait et al., 1998, Lundin, 1993)、ヒトの染色体上の相同領域の探索(e.g. Kasahara et al., 1996; Pebusque et al., 1998)などから示唆されている。様々な遺伝子族において、この時期にアイソフォームの多様化がみられることも、このことを示唆している。

アイソフォームの多様化は、無顎類と有顎類の分岐の前と後に見られた。このことから、遺伝子重複は無顎類と有顎類の分岐の前と後に起きたと考えられる。あるいは、遺伝子重複が起きた時期と無顎類と有顎類の分岐の時期が接近しているために、実際は無顎類と有顎類の分岐前(あるいは後)の時期に限定して起きているのに、解析からははっきりとした結果が得られていないという可能性もある。この時期の遺伝子重複が大規模なゲノムの重複によるならば、様々な遺伝子族で遺伝子重複の起きた時期に関して、同様の傾向がみられることが期待される。したがって今後、様々な遺伝子族について同様の解析を行い、データを蓄積することで、アイソフォームの多様化の起きた詳細な時期が明らかになるとと思われる。

表13 アイソフォームを作り出す遺伝子重複が起きた時期及びその回数

遺伝子族名 -サブタイプ名	無顎類と有顎類の分岐前 (回)	無顎類と有顎類の分岐後 (回)
PTK		
-FGFR	2.6	1.3
-Eph	6.5	2.3
-src	7.2	1.7
-PDGFR	3.9	0.8
Aldolase	1.3	1.2
Enolase	2.1	1.1
Complements	2.4	0.1

Suga et al., 1999bに基づいた。

PTP遺伝子族のアイソフォームが遺伝子重複により作り出された時期は、それ以外の時期と比べて分子進化速度が高かった。同様の傾向が、少なくとも節足動物と脊椎動物の分岐後、硬骨魚類と他の有顎類の分岐前にアイソフォームの多様化が起きた他の遺伝子族でもみられることが、岩部らにより示されている(Iwabe et al., 1996)。このことから、遺伝子重複がみられた時期はそれ以外の時期に比べて、機能的制約が弱かったことが示唆される。そのため、遺伝子重複の起きる頻度、あるいは作り出された重複遺伝子が集団に固定する機会、が多かったのかもしれない。限らの解析から、機能は同じだが発現する組織の異なる遺伝子間で分子進化速度が異なることが示され、組織レベルの機能的制約というもの存在が示唆されている(Kuma et al., 1995a)。このことから宮田らは、脊椎動物進化の初期に起きた発現する組織の異なるアイソフォームの多様化の背景に、組織レベルでの機能的制約の緩和があったのではないかと論じている(Miyata et al., 1994)。前節1-1の結果から、サブタイプの多様化が見られた多細胞動物進化の初期にも、分子進化速度の上昇が見られることがわかった。この時期にも同様に、何らかの機能的制約の緩和があったのかもしれない。

PTP遺伝子族のアイソフォームの多様化がみられた時期に起きた生物の進化

アイソフォームの多様化がみられた時期は、頭索動物と脊索動物の分岐後、軟骨魚類とその他の有顎類の分岐前であった。この時期はどのような時期だったのか。三胚葉動物は主に二つの系統、すなわち旧口動物と新口動物に分かれるが、そのうち新口動物の系統が、脊椎動物に至る進化を遂げた。ウニ・ヒトデなどの棘皮動物門やギボシムシの属する半索動物門の分岐後、脊索をもつ動物が現れる。原始的な脊索動物であるナメクジウオの属する頭索動物には背骨は認められず、脊椎動物のもつ多くの進化した構造も見られない。これに対して脊椎をもつ動物は、脳をはじめ感覚器・消化器・循環器・内分泌器・泌尿器系の様々な器官を進化させた。有顎類が出現した頃には、現在みられる脊椎動物の基本的な体制は完成していたと考えられている(Romer & Parsons, 1986)。アイソフォームの多様化のみられた時期は、この脊椎動物進化の初期に相当し、脊椎動物がその基本的体制を獲得した時期である。したがって、この時期にみられたアイソフォームの多様化と脊椎動物の初期進化との間には関連があるのかもしれない。

ヌタウナギが属する無顎類は、脊椎動物の中で最初に分岐したと考えられている。無顎類は脊椎や脳を持ち、頭索動物と比べれば複雑な体制をしている。しかし、顎や対鰭がなく、臍臓や胸腺・脾臓といった造血器官が独立した形で認められないなど、きわめて原始的な脊椎動物である(Romer & Parsons, 1986)。すなわち無顎類は脊椎動物が基本的な体制を確立するまでの、過渡期に分岐した生物なのかもしれない。アイソフォームの多様化と脊椎動物の初期進化との関連性を検証する上で、今後、ヌタウナギのPTP遺伝子の各アイソフォームが、どの組織で発現し、どのような機能を持っているかを解明することは重要である。例えば、PTPN6サブタイプのPTP1Cアイソフォームは、血球系細胞および胸腺・脾臓などの造血細胞に限定的な発現が認められる(Streuli, 1996)。PTP1Cアイソフォームに属する遺伝子に突然変異のみられるマウスは、自己免疫疾患等の症状を起こすことが知られている(Bignon & Siminovitch, 1994)。胸腺や脾臓といった器官の存在しないヌタウナギにおける、このアイソフォームとオーソログスな関係にあるhgPTPN6a遺伝子の発現場所や機能の解明は、脊椎動物の器官や組織の進化の分子基盤の理解につながると思われる。

また、アイソフォームの差異はしばしば発現する組織にみられた(表10参照)。しかしPTP遺伝子族については、主に哺乳類でしか発現する組織が同定されておらず、哺乳類以外の両生類・魚類などでの発現組織はわかっていない。哺乳類以外の脊椎動物でも、発現する組織は、哺乳類のホモログと共通しているのだろうか。上述のように、アイソフォームを作り出した遺伝子重複は、染色体レベルの大きな重複、あるいはゲノム全体の重複であったことが示唆されている。その場合には、アイソフォームを作り出した遺伝子重複

は、発現する組織を特定するプロモーター領域を伴う遺伝子重複であったといえる。もし、アイソフォームの発現組織が脊椎動物進化の初期には確立しており、すなわちプロモーター領域の機能（発現する組織の特定）が確立しており、かつ現在までその機能が保存されているならば、脊椎動物を通じて発現する組織が共通していることになる。そうでない場合には、脊椎動物の系統により発現する組織は異なっていると考えられる。発現する組織の確立した時期が、脊椎動物進化の初期であった場合には、アイソフォームの多様化と脊椎動物の組織の進化との関連を示唆することになるかもしれない。このことは、今後検証していく必要がある。

PTP遺伝子族のアイソフォームの多様化がみられない時期に起きた生物の進化

多様な動物が分岐したカンブリア爆発の起きた時期や、陸上脊椎動物が進化した時期などには、アイソフォームの多様化はみられなかった。アイソフォームの多様化がみられたのは、無顎類と有顎類の分岐の時期周辺（約5億年前、Dickerson 1971）であり、これはカンブリア爆発の起きた後である。したがって、特に多様な多細胞動物が進化したカンブリア爆発は、サブタイプの多様化ともアイソフォームの多様化とも、その時期のずれから考えて、関連の無いことが示唆される。すなわち、カンブリア爆発による多様な動物の出現は、新しい遺伝子を作り出すことなく達成されたと考えられる。このことについては、総合議論の章で詳しく考察する。

PTPR3サブタイプのアイソフォームの多様化

PTPR3サブタイプは、既存データの解析から、異なる傾向を示すことがわかった。図19はPTPR3サブタイプの分子系統樹である。アウトグループには他の受容体型サブタイプに属する遺伝子を用いた（網掛け）。この系統樹から、アイソフォームを作り出した遺伝子重複（菱形）の起きた時期が、海綿動物と真正後生動物の分岐（黒丸）後、節足動物と脊椎動物の分岐（白丸）前であることがわかる。この時期はまさにカンブリア爆発が起きた時期に相当する。ただし、カワカイメンのsPTPR3がパラログスな遺伝子であって、PTPR3サブタイプの各アイソフォームにより近縁な海綿のPTP遺伝子が存在し、遺伝子重複がサブタイプの多様化の見られた多細胞動物進化の初期に起きた可能性もある。特にPTPR3のアイソフォームの差異は、発現する組織というよりは、細胞外ドメインであるFNIIIドメインの数にみられ、異なる機能を果たしている可能性もある（表9参照）。このことを追求するためには、今後、原始的な多細胞動物である海綿動物、二胚葉動物である刺胞動物などの様々な動物から、PTPR3サブタイプに限定したPTP遺伝子の探査を行い、

分子進化学的解析を行う必要があるだろう。

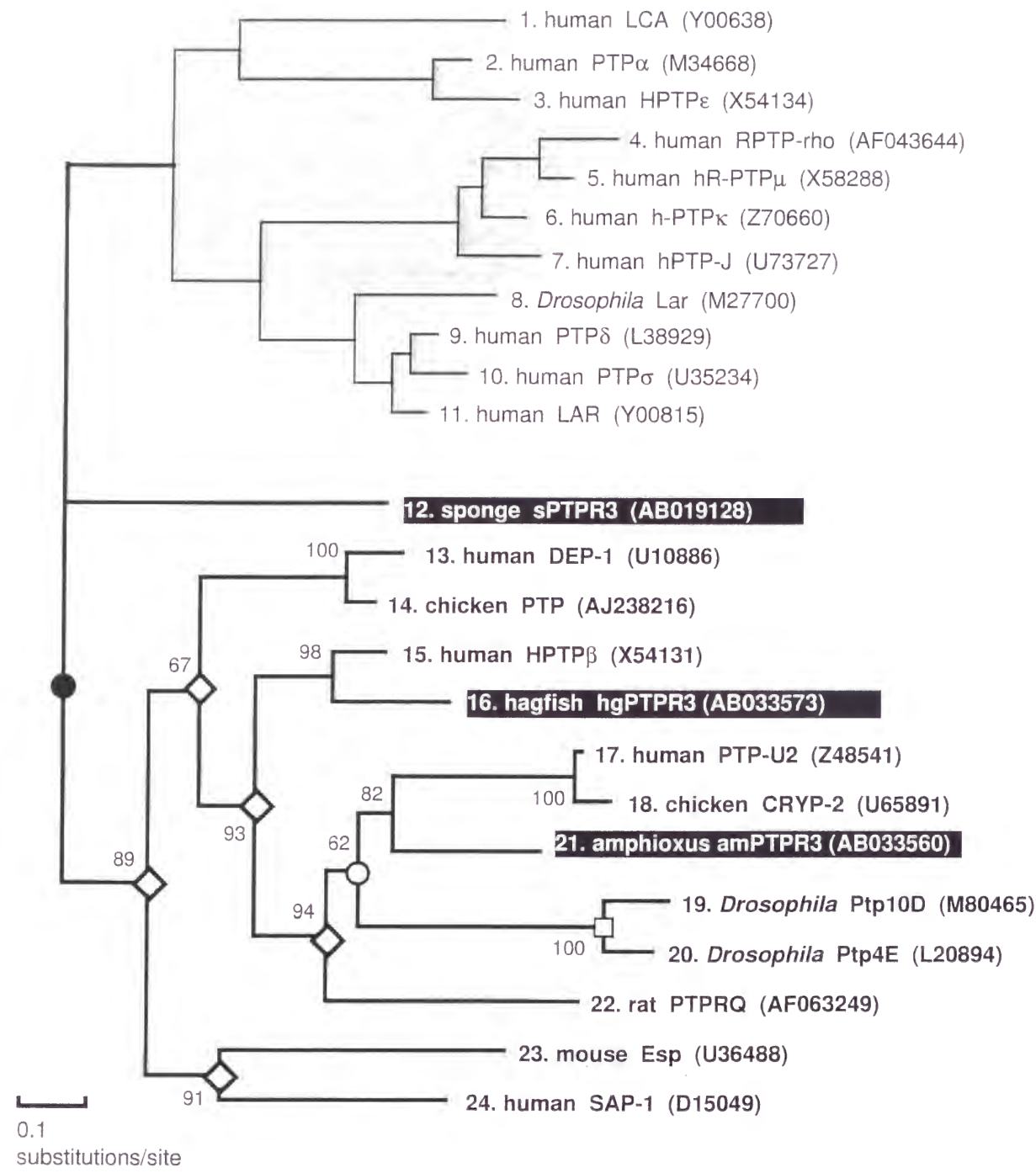


図19 PTPR3サブタイプの分子系統樹

PTPドメインの約3分の2に相当するアライメント領域（142アミノ酸座位）に基づき、発見的探索による最尤法で推定された最尤系統樹を示した。アライメントは付録5-8に示した。アウトグループには受容体型の他のサブタイプに属する遺伝子を用いた（網掛け）。遺伝子配列を決定した遺伝子は白抜き文字で示した。配列には生物種名、遺伝子名、GenBankのアクセッションナンバーを記載した。ノード付近に示した数字はLBPである。

黒丸は海綿動物と真正後生動物の分岐を、白丸は節足動物と脊椎動物の分岐を示す。菱形は脊椎動物のアイソフォームを作り出した遺伝子重複を、四角は脊椎動物以外の系統で起きた遺伝子重複を表す。

2. プロテインセリン・スレオニンホスファターゼ(PP)遺伝子族

2-1. PP遺伝子族のサブタイプの多様化

プロテインセリン・スレオニンホスファターゼ(PP)は、プロテインチロシンホスファターゼ(PTP)と異なり、古細菌や真核生物に幅広く存在している。このうち真核生物には、複数種類のPP遺伝子が知られている(Shenolikar, 1994; Mumby & Walter, 1993 for review)。古細菌のPP遺伝子は一種類しか知られていない(e.g. Leng et al., 1995)ことから、真核生物に見られる複数種類のPP遺伝子は、古細菌と真核生物の分岐後、真核生物の系統で独自に作り出されたものと考えられる。これらの真核生物のPP遺伝子は、調節サブユニットの種類の違いや、PP遺伝子のドメイン構造の違いから、表14のように分類されている。ここでいう調節サブユニットとは、PP遺伝子と異なる遺伝子にコードされ、PP遺伝子と複合体を形成して働くタンパク質のことである。この調節サブユニットやドメイン構造は、基質特異性、調節を受ける分子、細胞内局在の決定等に関与している(Hubbard & Cohen, 1993; Faux & Scott, 1996 for review)。表14で分類された遺伝子群、すなわち同じ調節サブユニットと複合体を形成するPP遺伝子群、あるいは同じドメイン構造を持つPP遺伝子群は、共通の機能を担うまとまったグループと考えられる。そこで、本論文ではこれをサブタイプと呼ぶことにする。

真核生物でみられるPP遺伝子の多様なサブタイプは、真核生物の進化の過程においていつ頃作り出されたのだろうか。この問題を明らかにする目的で、ギアルディアという生物に注目した。ギアルディアは、ミトコンドリアを持たず、細胞内小器官もあまり発達していない寄生性の原生生物である(Gillin et al., 1996 for review)。rRNA及びタンパク質を用いた分子系統解析から、真核生物の中で最初に他の真核生物と分岐した原生生物だと考えられている(図20、Sogin et al., 1989; Hashimoto et al., 1994)。このギアルディアが分岐する時点では、どれ程のサブタイプが作り出されていただろうか。現在ギアルディアは、その進化的及び医学的重要性から全塩基配列を決定するゲノムプロジェクトが進行中である(Smith et al., 1998)。そこで、このゲノムプロジェクトで公開されているデータベースを利用し、PP遺伝子族のサブタイプが遺伝子重複により作り出された時期を明らかにする目的で、分子進化学的解析を行った。

表14 プロテインセリン・スレオニンホスファターゼ遺伝子の分類

サブタイプ	調節サブユニット	ヒトの主な遺伝子
PP1	I-1, I-2, G, M, N	PP1α, PP1β, PP1γ
PP2A	A, B', B''	PP2Aα, PP2Aβ
PP2B	B	PP2Bα, PP2Bβ, PP2Bγ

サブタイプ	ドメイン構造	ヒト主な遺伝子
PP5	Tetratricopeptide repeats	PP5
rdgC	EF hand motifs	PPEF

Andreeva & Kutuzov, 1999; Shenolikar, 1994; Mumby & Walter, 1993 に基づいた。

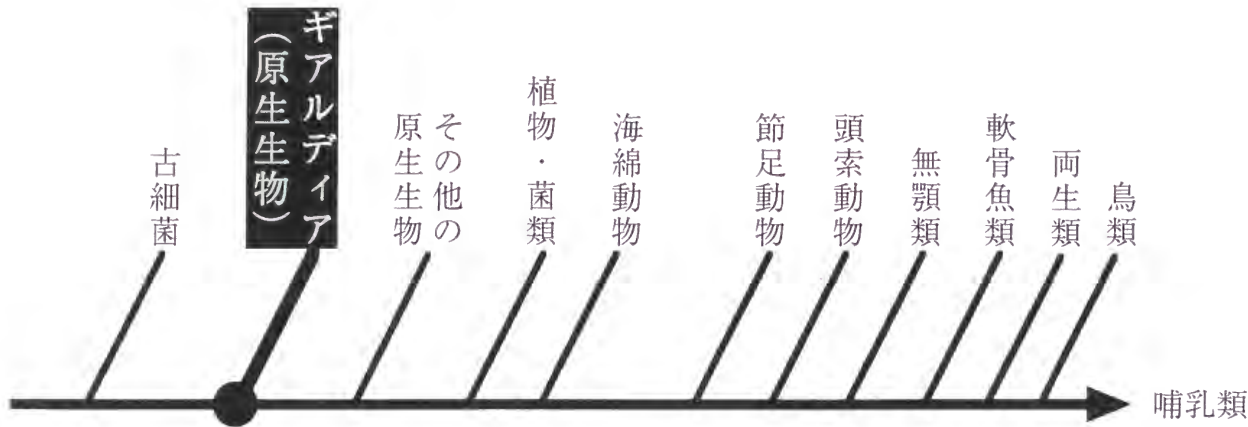


図20 ギアルディアの系統的位置

本節で注目する原生生物ギアルディアの系統的位置を模式的に示した。枝の長さは概ね時間を反映しているが正確ではない。

結果

ギアルディアのPP遺伝子

ギアルディアのDNA配列データベースを対象に、PP遺伝子の相同性検索を行った。このデータベースは28328本のゲノムDNA断片配列から成り、断片配列の平均長は884ヌクレオチドである。本論文では、材料と方法の章に記した手順で断片配列を連結し再構築したデータベースに対して相同性検索を行った。

その結果、13種類の配列にPP遺伝子との相同性が認められた。13種類の配列は、本論文において仮に、遺伝子名をgPP01～gPP13と名付けた。

gPP01～gPP09の9種類の遺伝子は、PPドメインのほぼ全領域を含み、他の真核生物のPP遺伝子とは30%～81%、古細菌のPP遺伝子とは26%～42%の一致度を示した。また、PP遺伝子族で保存しているモチーフ 'DXH-GDXXD-GNHD/E' (Zhuo et al., 1994)が保存していた(図21)。したがってこの9遺伝子は、ギアルディアにおいてPPとして機能しているものと思われる(各遺伝子を構成するゲノムDNA断片の配列番号を付録4に示した)。gPP10、gPP11、gPP12は断片配列であった。またgPP13は他の真核生物のPP遺伝子と23%～32%、古細菌のPP遺伝子とは20%～28%と一致度が低かった。これらgPP10～13の4遺伝子は、分子系統樹推定の信頼性を低下させる可能性があるため、以降の解析では除外した。


```

1. human PPPICC (X74008)
2. Giardia gpp01
3. Giardia gpp02
4. Giardia gpp03
5. Giardia gpp04
6. Giardia gpp05
7. Giardia gpp06
8. Giardia gpp07
9. Giardia gpp08
10. Giardia gpp09

1. ( 58) LKICGDIHGQYYDLLRLFEYGGFPPESEN-----
2. ( 67) IKICGDIHGQFYDLLRLFEYGGFPATN-----
3. ( 52) ITLVGDIHGQFYDLREMF SIGGSIPDVQ-----
4. (116) IYIVGDIHGQLHDILNYFNVLGSP LTKN-----
5. ( 95) LNVCGDVHGQFYDVRYMF EFGGKIVD GAR-----
6. (257) LAIVGDIHGSLPDLCAINEKLVPM MFSEKMK-----
7. ( 94) VTVVGDIHGQYHDLIEMFRICGRSPYTN-----
8. ( 39) IAVIGDLHGDAESTKYIFQKVLNKT FME SGM-----
9. ( 35) LVVVGDIHGTFEVI GKILATEGFPTDNE IV-----
10. ( 91) VNICGDLHGQFYDVVKLFDVGGRIDNSTPXGGKHDRV PANLKPNTSMS SNAKALRVHPTP
      ** **

1. --YLF L G D Y V D R G K Q S L E T I C L L L A Y K I K Y P E N F F L L R G N H E C A S I (193)
2. --Y I F L G D Y V D R G K Q S L E C I I L L L C Y K I K Y P E N F F M L R G N H E C S S I (199)
3. --Y L F L G D Y V D R G Y F S V E T F L L L I A F K V K F P T R I S L L R G N H E S R Q I (190)
4. --Y L F L G D V V D R G Y Q S V E C L L L M L V C K I K R P N N F F I R G N H E T D L (205)
5. --Y L F L G D Y V D R G Y S L E T I T L L L C L K V N Y P A R I Y L L R G N H E C R A V (179)
6. --A V F L G D Y V D R G P K G H T V V T A L L C L K L C F P D R V F L L R G N H E T V S M (187)
7. --Y L F M G D Y V D R G Y S V E C I T L L V L L K I K Y R N R V T M L R G N H E S R Q V (186)
8. --A V F L G D Y V D R G Q H G I N V L T S I L A A K V V Y G D R V I T I R G N H E S E N L (231)
9. --Y L F N G D Y V D R G S F S I E I F I V L L A I K I A R P D A V H L L R G N H E T E S V (234)
10. K K Y L F L G D Y V D R G Y F S L E T I T L L Y L L K L M Y P T Q I Y L I R G N H E C R S I (188)
      *  **  *****

```

図21 ギアルディアのPP遺伝子のアライメント

相同性検索の結果、PP遺伝子と相同性の認められたギアルディアの遺伝子のアライメント。比較のためヒトのPPPICC遺伝子を最上部に示した。PP遺伝子族で保存しているモチーフ 'DXH-GDXXD-GNHD/E' (Zhuo et al., 1994)を白抜き文字で示した。-はギャップを表す。*は全ての配列で保存しているアミノ酸を示す。括弧内はその領域に存在するアミノ酸の総数。

PP遺伝子族の分子系統樹の推定

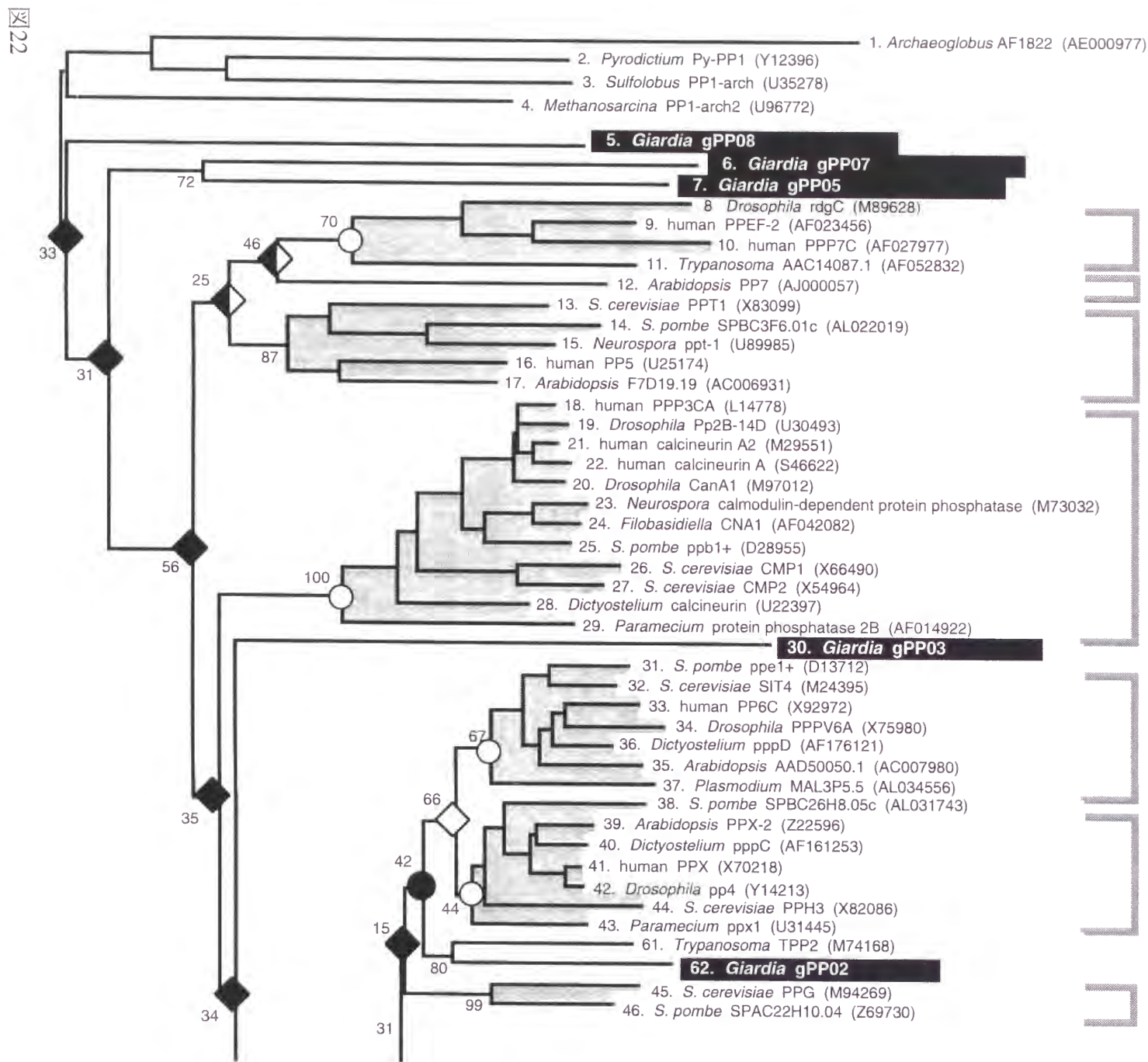
PP遺伝子族の多様なサブタイプが遺伝子重複により作り出された時期を知る目的で、既存のデータベース中に存在するPP遺伝子とギアルディアのPP遺伝子を併せて、近隣結合法により分子系統樹を推定した（図22）。アウトグループには、古細菌のPP遺伝子群を用いた。

その結果、ギアルディアの遺伝子（白抜き文字で示した）gPP06はPP2Aサブタイプに分類された。PP2AサブタイプのC末端には、メチルエステル化に必要なアミノ酸-YFLが保存している(Lee et al., 1996)が、gPP06にもこのアミノ酸が保存していた。このことはgPP06がPP2Aサブタイプに属することを支持する。gPP01はPP1サブタイプ及びPPYサブタイプと、gPP02はPP4サブタイプ及びPP6サブタイプとそれぞれオーソログな関係にある遺伝子である。残りの6種類の遺伝子は、既知のサブタイプに属さなかった。

PP遺伝子族のサブタイプを作り出す遺伝子重複が起きた時期の推定

図22の分子系統樹から、PP遺伝子の多様なサブタイプが作り出された時期が推定できる。分子系統樹上で、異なるサブタイプを作り出した遺伝子重複に菱形の印を付した。そのような遺伝子重複16回のうち12回が、真核生物最古の種分岐であるギアルディアとその他の真核生物の分岐（黒丸）前に起きていた（黒の菱形）。ギアルディアの分岐後に起きた遺伝子重複は2回である（白の菱形）。残りの2回の遺伝子重複については、遺伝子重複の起きた時期がギアルディアの分岐前なのか後なのか明らかでない（白黒の菱形）。しかし、少なくともゾウリムシ・マラリヤ原虫といった原生生物と、他の真核生物との分岐（白丸）前である。

この結果は、図22の分子系統樹に基づいて得られた結果である。図22の分子系統樹には、分岐点のブートストラップ確率が低く、信頼性の低い分岐も含まれる。そこで、異なるサブタイプを作り出した遺伝子重複が起きた時期を統計的に検定する目的で、ブートストラップ検定を行った（図23）。その結果、異なるサブタイプを作り出した遺伝子重複のうち、古細菌と真核生物の分岐後、ギアルディアの分岐前に起きたものは平均で10.9回、ギアルディアの分岐後に起きたものは平均で1.4回、ギアルディアの分岐前なのか後なのか不明のものが平均で1.0回となった。この値は図22の分子系統樹に基づいて得られた値と概ね一致する。このことから、異なるサブタイプを作り出した遺伝子重複のほとんどが、真核生物最古の種分岐であるギアルディアとその他の真核生物の分岐前に起きていたという結果が確かめられた。



archaeobacteria

rdgC



EF hand motifs

PP7

PP5



Tetratricopeptide repeats

PP2B

PP6

PP4

PPG

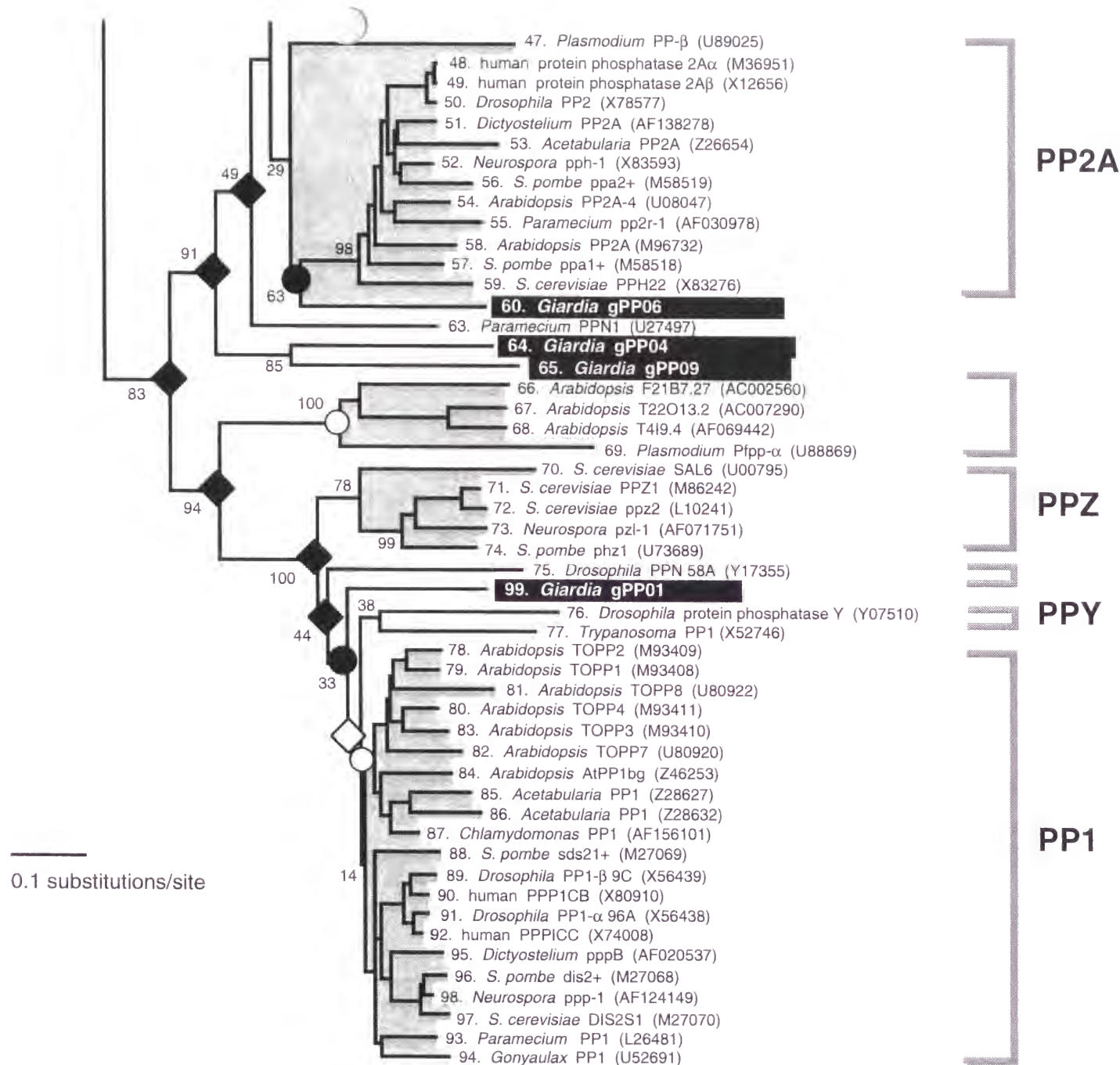


図22 PP遺伝子族の分子系統樹

PPドメインの約3分の2に相当するアライメント領域（182アミノ酸座位）に基づき、近隣結合法により推定した分子系統樹を示した。アライメントは付録5-9に示した。アウトグループには古細菌のPP遺伝子群を用いた。ギアルディアの遺伝子は白抜き文字で示した。配列には生物種名、遺伝子名、GenBankのアクセッションナンバーを記載した。ドメイン構造が同じ遺伝子群（PP5, rdgCサブタイプ、ドメイン構造をCohen, 1997に基づき系統樹の右側に示した）、調節サブユニットが同じであると考えられる遺伝子群（PP1, PP2A, PP2Bサブタイプ）、遺伝子内にPPドメイン以外に相同性のある領域がある遺伝子群（66番～69番のシロイヌナズナとマラリヤ原虫の遺伝子）を同じサブタイプに属する遺伝子と考え網掛けで示し、サブタイプ名を右側に記した。それ以外の遺伝子群についてはAndreeva & Kutuzov, 1999; Shenolikar, 1994; Mumby & Walter, 1993を参考にしてサブタイプと定義した。サブタイプ名の存在しないものは記していない。脊椎動物からはヒトの遺伝子、無脊椎動物からはショウジョウバエの遺伝子、高等植物からはシロイヌナズナの遺伝子を代表させて系統樹推定に用いた。比較的最近（同義置換数が飽和していない程度）に分岐した同一生物種の遺伝子については一方を代表させた。ノード付近に示した数字はブートストラップ確率を示す。

黒丸はギアルディアとその他の真核生物の分岐を、白丸はギアルディア以外の原生生物とその他の真核生物の分岐を表す。菱形は異なるサブタイプを作り出した遺伝子重複を示す。このうち黒い菱形は古細菌と真核生物の分岐後ギアルディアの分岐前に起きた遺伝子重複、白の菱形はギアルディアの分岐後に起きた遺伝子重複、白黒の菱形はギアルディアの分岐前か後か不明の遺伝子重複である。

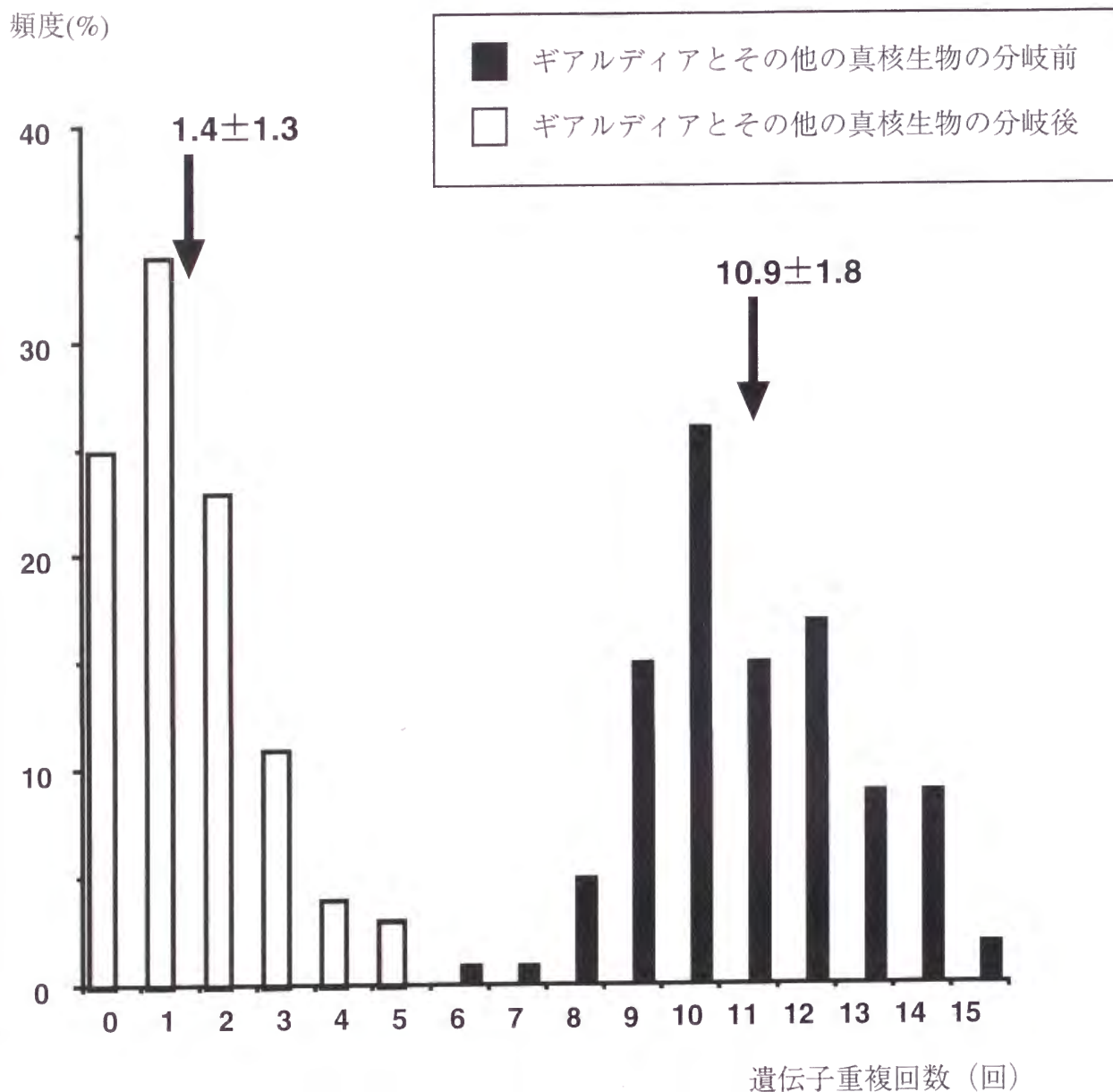


図23 PP遺伝子族のサブタイプを作り出す遺伝子重複が起きた時期のブートストラップ検定

ブートストラップ検定に基づき、異なるサブタイプを作り出す遺伝子重複が起きた時期及びその回数の頻度分布を求めヒストグラムに示した。矢印は平均を示す。異なるサブタイプを作り出す遺伝子重複のうち、古細菌と真核生物の分岐後、ギアルディアとその他の真核生物の分岐前に起きたものを黒、ギアルディアとその他の真核生物の分岐後に起きたものを白で表した。ギアルディア以外の原生生物の遺伝子は、分岐の順序が明確でなく、またその分岐がギアルディアと接近する場合もあるため、ギアルディアの分岐前に起きた遺伝子重複回数を過剰に見積もる危険性がある。そこでこのブートストラップ検定の解析は、図22の遺伝子配列からギアルディア以外の原生生物の配列を除外して行った。

考察

他の遺伝子族との比較

PP遺伝子族の多様なサブタイプの多くが遺伝子重複により作り出された時期は、古細菌と真核生物分岐後、ギアルディアと他の真核生物の分岐前であった(図24)。また、ギアルディアとの分岐後には、新しいサブタイプはほとんど作り出されていなかった。同様の結果が、細胞内の小胞輸送に関わるrab遺伝子族や、細胞骨格タンパク質であるキネシン遺伝子族でも見られることが報告されている(Iwabe et al., manuscript in preparation)。

PTP遺伝子族と比べてPP遺伝子族のサブタイプの多様化が起きた時期は、非常に古い時期であった。このことはPP遺伝子とPTP遺伝子の役割の違いに関連すると考えられる。PTP遺伝子は、動物において、細胞間コミュニケーションや細胞接着などで重要な役割を果たす(Charbonneau & Tonks, 1992; Hunter, 1995; Streuli, 1996 for review)。これに対してPP遺伝子は、主に転写や細胞周期の制御など真核生物全般に重要だと考えられる役割を果たしている(Shenolikar, 1994; Mumby & Walter, 1993 for review)。このような役割の違いが、PTP遺伝子族とPP遺伝子族の多様化の起きた時期の違いを反映していると思われる。

また、同様の傾向はリン酸化を行う酵素であるプロテインキナーゼにもみられる。すなわち、プロテインチロシンキナーゼ(PTK)遺伝子族では、多細胞動物進化の初期と脊椎動物進化の初期に遺伝子重複による多様化がみられる(Suga et al., 1999a; 1999b)。これに対してプロテインセリン・スレオニンキナーゼ(PSK)遺伝子族では、真核生物進化の初期、多細胞動物進化の初期、脊椎動物進化の初期のそれぞれの時期に遺伝子重複による多様化がみられることが示唆されている(小野, 1997)。したがって、チロシン残基のリン酸化・脱リン酸化を行うPTK遺伝子族とPTP遺伝子族、及びセリン・スレオニン残基のリン酸化・脱リン酸化を行うPSK遺伝子族とPP遺伝子族、の多様化の時期は一致している。このように、リン酸化・脱リン酸化によるシグナル伝達系は、リン酸化を受けるアミノ酸残基の違いに応じて、異なる時期に多様化し、両者を使い分けて進化したと考えられる。

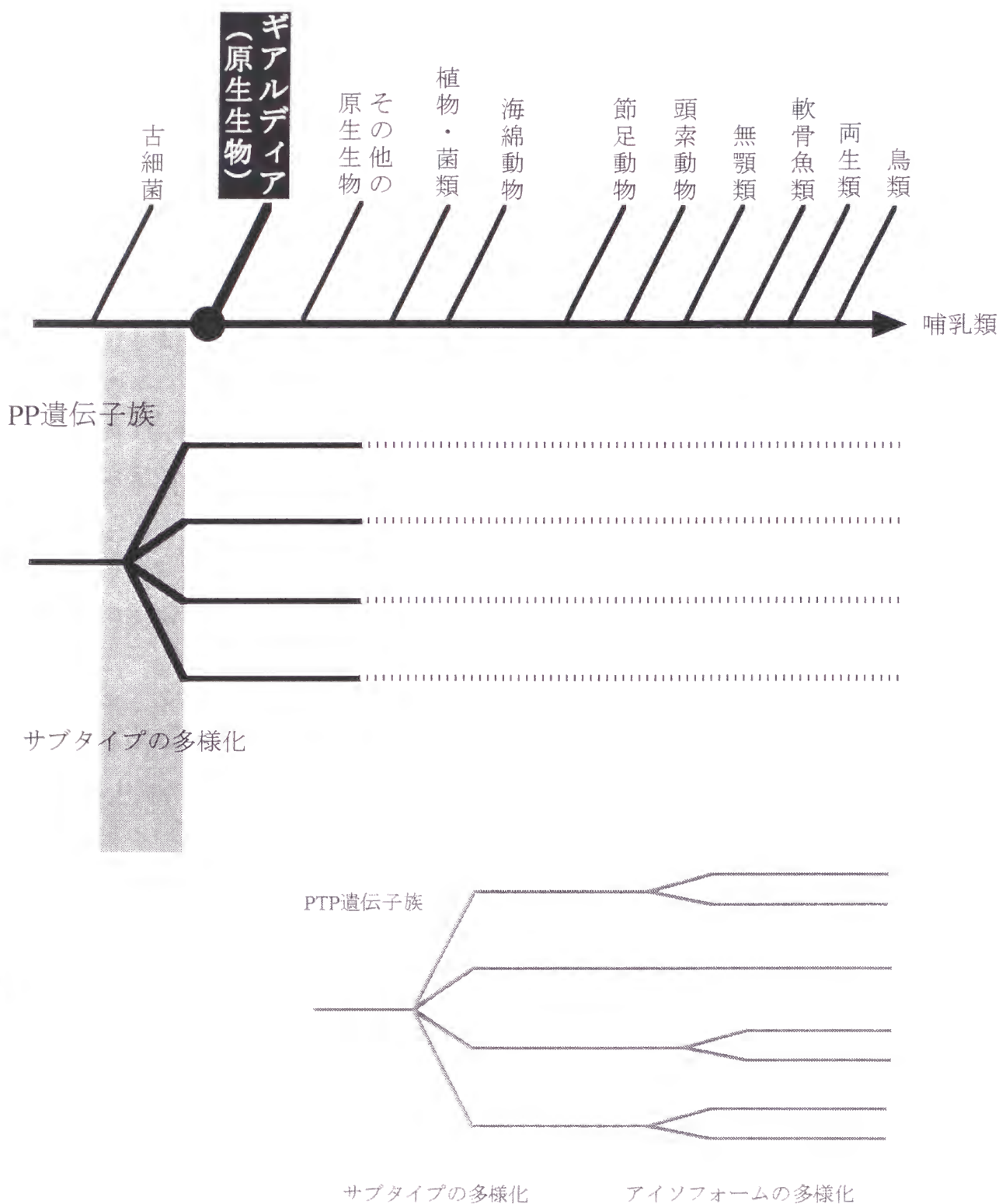


図24 PP遺伝子族のサブタイプの多様化のまとめ

PP遺伝子族のサブタイプの多様化の時期を模式的に示した。本論文の結果から明らかとなった、サブタイプの多様化が見られた時期を網掛けで示した。枝の長さは概ね時間を反映しているが正確ではない。比較のため、下にPTP遺伝子族の多様化が見られた時期（前節1-1及び1-2参照）を示した。

PP遺伝子の調節サブユニット及びドメイン構造の起源

PP遺伝子の機能の異なるサブタイプを作り出した遺伝子重複の多くは、古細菌と真核生物の分岐後、ギアルディアの分岐前に起きていた。それでは、各サブタイプを特徴づける調節サブユニットやドメイン構造の起源はいつだろうか。調節サブユニットやドメイン構造がギアルディアにも存在し、その調節サブユニットが既知の組み合わせで複合体を形成し機能していた場合には、PP遺伝子族のサブタイプの機能の確立もまた、ギアルディアの分岐前という古い時期に完了していたことを示唆する。

各調節サブユニットを問い合わせ配列として、ギアルディアのデータベースに対して相同性検索を行った。その結果、PP2Aサブタイプの調節サブユニットと相同性のある配列がギアルディアにも存在した。これらの遺伝子がギアルディアにおいても、PP2Aサブタイプと複合体を形成するかを調べることは、調節サブユニットの起源を知る上で重要である。また現時点では、PP1サブタイプやPP2Bサブタイプの調節サブユニットと相同性のあるギアルディアの配列は存在しなかった。これらの調節サブユニットは、ギアルディアの分岐後に使われるようになったのかもしれない。ギアルディアの全塩基配列が決定すればこのことが検証される。

PP5サブタイプやrdgCサブタイプは特徴的なドメインを持つサブタイプである。現時点では、これらのサブタイプのホモログとなるギアルディアの遺伝子は見つかっていない。ただしギアルディアのgPP05のN末端側にはTetratricopeptide Repeatsが存在した（図25）。このドメイン構造はPP5サブタイプに特徴的なものである。gPP05は、図22の分子系統樹からはPP5サブタイプには分類されなかった。しかしこれは、ギアルディア遺伝子配列内のエラーに起因するlong-branch attractionのアーティファクトで（後述）、実際はPP5サブタイプに属する、という可能性も残されている。その場合には、PP5サブタイプを作り出したドメインシャッフリングは、ギアルディアの分岐前に起きたことになる。

1. *Giardia* gPP05
2. *Saccharomyces cerevisiae* PPT1 (X83099)
3. *Schizosaccharomyces pombe* SPBC3F6.01c (AL022019)
4. *Neurospora crassa* ppt-1 (U89985)
5. *Homo sapiens* PP5 (U25174)
6. *Arabidopsis thaliana* F7D19.19 (AC006931)

```

1. ( 7) ADELKALANEAFRAGHYPRAVELYTQAIEQLDSPRAEYFTNRSAHFQNGDFVAAKADAEAAIDID
2. (11) ALERKNEGNVFEKHFLEKFLKAEKYTEAID-LDSTQSIYFSNRAFAHFKVDNFQSALNDCDEAIKLD
3. ( 4) ALELKNEANKFLKEGHIVQAIDLYTKAIE-LDSTNAILYSNRSLAHLKSEDYGLAINDASKAIECD
4. ( 7) AIAFKNEGNKAFAAHDWPKAIEFYDKAIE-LNDKEPTFWSNRAQHLKTEAYGYAIRDATKAIELN
5. (19) AEELKTQANDYFKAKDYENAIKFYSQAIE-LNPSNAIYYGNRSLAYLRTECYGYALGDATRAIELD
6. (12) AEEFKSQANEAFKGHKYSSAIDLYTKAIE-LNSNNAVYWANRAFAHTKLEEYGSIAIQDASKAIEVD
      *  *  *                *  *  *  *  *                *  *  *  *  *

```

```

1. PNSTKGYWRIAECERIMHQLDKSIEYYRRALETKPNNRKIVLAIRDVQRLIDIAPKPSHGL (386)
2. PKNIKAYHRRALSCMALLEFKKARKDLNVLLKAKPNDPAATKALLTCDRFIREERFRKAI (377)
3. PEYAKAYFRRATAHIAIFQPKAEAVGDFRKALALAPSDPAARKKLRECEQLVKRIRFQEAI (344)
4. PGFVKAYYRRATAYAAILNPKEAVKDFKTCVKIAPDNKDAKLKLVCECEKIVRQLAEFFAAI (347)
5. KKYIKGYRRAASNMA LGKFRAALRDYETVVVKVPHDKDAKMKYQECNKIVKQKAFFERAI (347)
6. SRYSGYYRRGAAYLAMGKFKDALKDFQQVKRLSPNDPDATRKLKECEKAVMKLKFEETAI (396)
      *  *  *                                *

```

図25 Tetratricopeptide Repeatsのアライメント

ギアルディアのgPP05遺伝子とPP5サブタイプに属する遺伝子の、Tetratricopeptide Repeats領域のアライメント。-はギャップを表す。*は全ての配列で保存しているアミノ酸を示す。括弧内はその領域に存在するアミノ酸の総数。

PP遺伝子族のサブタイプの多様化がみられた時期に起きた生物の進化

PP遺伝子族のサブタイプの多様化は、古細菌と真核生物の分岐後、ギアルディアとその他の真核生物の分岐前に起きていた。この時期はどのような時期だったのか。地球が誕生したのは、今から約46億年前だと推定される(Futuyma, 1998)。その地球における最古の生命の化石は34~31億年前のもので、その形態は現在のシアノバクテリアに類似している(Futuyma, 1998; Margulis & Schwartz, 1998)。その後約10億年間は、核をもたない単細胞原核生物（真正細菌及び古細菌）の時代が続いたが、約21億年前になると最古の真核生物が現れる(Han & Runnegar, 1992)。核を持つ真核生物は、原核生物と比べて細胞サイズやゲノムサイズが大きく、細胞内小器官や細胞骨格をもつなど細胞内構造がより複雑となった(Alberts et al., 1994)。ギアルディアはこの真核生物の中で最初に分岐した現存の原生生物と考えられている(Sogin et al., 1989; Hashimoto et al., 1994)。PP遺伝子族でサブタイプの多様化が見られた時期は、この真核生物進化の初期に相当する。したがって、この時期にみられた遺伝子族のサブタイプの多様化と、真核生物の初期進化との間には何らかの関連性のある可能性がある。また、細胞内の小胞輸送に関わる遺伝子族や、細胞骨格タンパク質の遺伝子族でもこの時期にサブタイプの多様化がみられた(Iwabe et al., manuscript in preparation)。これらの遺伝子族の多様化が、真核生物のもつ細胞内構造の複雑化に寄与したのかもしれない。

ギアルディアは細胞内小器官が発達していない原生生物である(Gillin et al., 1996 for review)。かつては、ミトコンドリアをもたないことから、ミトコンドリア共生以前の原始的な生物「アーケゾア」とされていた(Cavalier-Smith, 1989)。しかし、最近の研究から、ミトコンドリアをかつては保持しており、二次的に失ったことが明らかとなった(Roger et al., 1998; Hashimoto et al., 1998)。本論文の解析から、ギアルディアにもPP遺伝子のサブタイプのほぼ完全なセットが存在することが示唆された。この結果も、ギアルディアの原始性が二次的な（おそらく寄生という生活環境に起因する）結果であることを意味するのかもしれない。

PP遺伝子族のサブタイプの多様化がみられない時期に起きた生物の進化

ギアルディアの分岐後も、多様な原生生物が出現したと考えられる(Cavalier-Smith, 1993)。このような時期には、PP遺伝子族の新しいサブタイプは作り出されていなかった。したがってこの時期に起きた生物の進化は、その時期のずれから考えて、PP遺伝子族のサブタイプの多様化とは関連のないことが示唆される。ただし、様々な原生生物の分岐年代等は不明な点が多く、多様な原生生物の進化と遺伝子重複の関連について検証する

ことは、今後の課題である。

サブタイプの多様化の後、動物の系統では各サブタイプで遺伝子重複がみられ、アイソフォームの多様化が起きている。これらの遺伝子重複と動物の進化との関連性については、次の節で検討する（2-2節）。

ギアルディアの遺伝子配列の問題点

ギアルディアのゲノムプロジェクトは未だ完了しておらず、ギアルディアのDNA配列データベースは予備的に公開されたものである。したがって配列内にエラーが含まれる可能性が指摘されている(http://evol3.mbl.edu/GiardiaHTML/giardia_data.html)。本解析では、材料と方法の章に記した手順により最終配列を決定することで、配列内のエラーが除外されるよう考慮したが、それでもエラーが含まれる可能性は残る。エラーを含んだギアルディアの遺伝子は、分子進化速度の高い配列と同様の振舞いを見せる可能性がある。一般に分子系統樹推定において、分子進化速度の高い配列は、本来の分岐位置よりも古い時期に分岐するよう誤って推定される傾向がある(long branch attraction, Felsenstein, 1978)。したがって、ギアルディアの遺伝子は、実際よりも古い時期に分岐すると誤って推定された可能性がある。

図22の分子系統樹において、gPP03～gPP09は、既存のサブタイプに属しなかった。しかし、これがlong branch attractionによるアーティファクトであり、実際には既存のサブタイプに属する遺伝子である可能性もある。

ギアルディアの分岐位置が実際よりも古くに推定されることは、ここで得られた結論に不利に働く。したがって、サブタイプを作り出した多くの遺伝子重複がギアルディアの分岐前に起きた、という結論に影響はない。

出芽酵母のPP遺伝子

PP遺伝子の13種類のサブタイプは、真核生物進化の初期に作り出したことが明らかとなった。全ゲノムの塩基配列が決定した出芽酵母(*Saccharomyces cerevisiae*, Clayton et al., 1997)は、この後に分岐する生物である。したがって、13種類全てのサブタイプを保持していることが期待される。ところが、細胞内において重要な役割を果たすと考えられるPP1、PP2A、PP2Bなどの基本的なサブタイプに属する遺伝子は保持しているものの、rdgC、PP-Yサブタイプなど5種類のサブタイプに属する遺伝子が存在しなかった（図22参照）。これらのサブタイプに属する遺伝子は、出芽酵母の系統において欠落したものと考えられる。このことは、出芽酵母の系統に遺伝子数を減らしてゲノムを小さくするような

制約が存在したことを示しているのかもしれない。あるいは、重複遺伝子は全てが必須ではなく、機能的に補完できるものが存在するために、系統によっては偶然に重複遺伝子の欠落が起きるのかもしれない。

2-2. PP遺伝子族のアイソフォームの多様化

真核生物はその進化の初期の段階で、13種類の機能の異なるPP遺伝子族のサブタイプを作り出していた。そして、それ以降新しいサブタイプはほとんど作り出されなかった。一方、いくつかのサブタイプは、その後遺伝子重複により遺伝子を増やしている。このうち本節では、動物の系統で起きた遺伝子重複に注目する。同じサブタイプに属する遺伝子同士は、複合体を形成する調節サブユニットやドメイン構造が同じであることから、基本的な機能は共通していると考えられる。遺伝子間には、明確な差異の知られていないものもあるが、PTP遺伝子族同様発現する組織が異なるものもある(Shenolikar, 1994; Mumby & Walter, 1993 for review)。本論文ではこれらをアイソフォームと呼ぶことにする。

これらのアイソフォームは、動物の系統において、遺伝子重複により作り出されたものである。それでは、PP遺伝子族のアイソフォームが作り出された時期はいつだろうか。前章（1-1及び1-2）で解析したPTP遺伝子族では、サブタイプの多様化が多細胞動物進化の初期、アイソフォームの多様化が脊椎動物進化の初期、の限定した時期に起きていた。PP遺伝子族でアイソフォームの多様化が起きた時期は、これらの時期に対応するのだろうか。このことを明らかにする目的で、多細胞動物の中で最初に分岐した海綿動物のカワカイメン、脊椎動物に最も近縁な無脊椎動物である頭索動物のナメクジウオ、脊椎動物の中で最初に分岐した無顎類のヌタウナギ、有顎類の中で最初に分岐した軟骨魚類のエイから、PP遺伝子の探査、及び分子進化学的解析を行った（図26）。

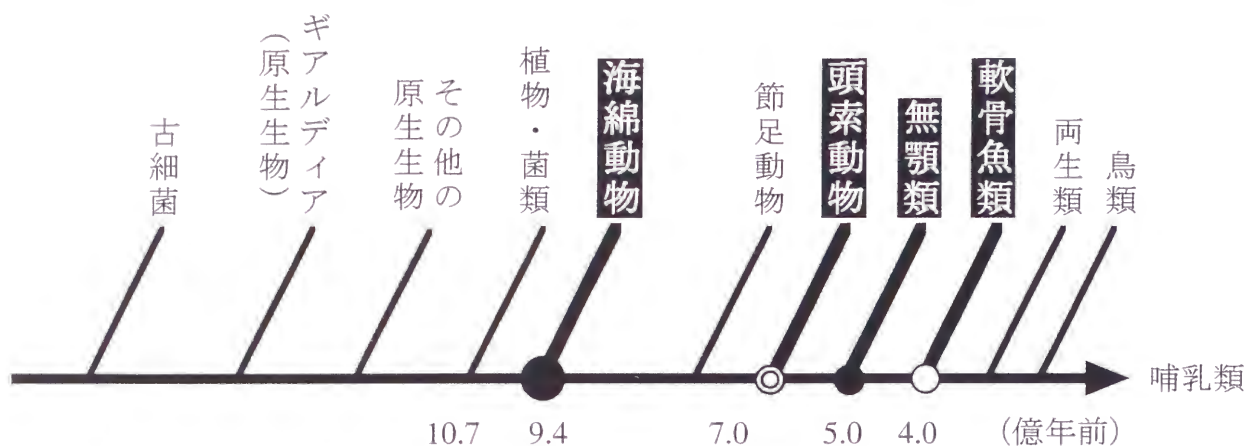


図26 海綿動物、頭索動物、無顎類、軟骨魚類の系統的位置

本節で注目する海綿動物、頭索動物、無顎類、軟骨魚類の系統的位置を模式的に示した図。枝の長さは概ね時間を反映しているが正確ではない。数字は分岐年代を示す(Dickerson, 1971; Dayhoff, 1978; Nikoh et al., 1997)。

結果

カワカイメン・ナメクジウオ・ヌタウナギ・エイのPP遺伝子のクローニング

RT-PCRによりカワカイメン・ナメクジウオ・ヌタウナギ・エイからPP遺伝子の探査を行った。その結果、カワカイメンから6種類、ナメクジウオから4種類、ヌタウナギから5種類、エイから9種類の、計24種類のPP遺伝子の単離に成功した（塩基配列及び推定アミノ酸配列は付録2-3に示した）。塩基配列を決定した領域は、PPドメインのN末側約1/3から終止コドンまでに相当する。

カワカイメンのPP遺伝子は他の真核生物のPP遺伝子と32～95%、ナメクジウオのPP遺伝子は32～100%、ヌタウナギのPP遺伝子は32～100.0%、エイのPP遺伝子は33～100%、の一致度を示した。また、単離した全てのPP遺伝子に、PP遺伝子族で保存しているモチーフ 'DXH-GDXXD-GNHD/E' (Zhuo et al., 1994)が保存していた（付録2-3参照）。したがってこれらの遺伝子は、PPとして機能しているものと思われる。

単離した遺伝子のサブタイプの同定

単離したカワカイメン・ナメクジウオ・ヌタウナギ・エイのPP遺伝子が、どのサブタイプに属するかを調べる目的で、既存のデータベース中に存在するPP遺伝子と、単離されたカワカイメン・ナメクジウオ・ヌタウナギ・エイのPP遺伝子を併せて、PP遺伝子族全体の分子系統樹を推定した（図27）。

その結果、単離した24種類の遺伝子（白抜き文字で示した）は全て既知のサブタイプに分類された。それぞれの遺伝子と各サブタイプに属する遺伝子は、高いブートストラップ確率で単系統群を形成しており、この分類が確実であることを示している。

PP1、PP2A、PP2Bの3つのサブタイプにおいては、アイソフォームの多様化がみられ、なおかつ、カワカイメン・ナメクジウオ・ヌタウナギ・エイからPP遺伝子が単離されている。そこでこの3つのサブタイプについて、アイソフォームを作り出した遺伝子重複が起きた時期の詳細な解析を行った。

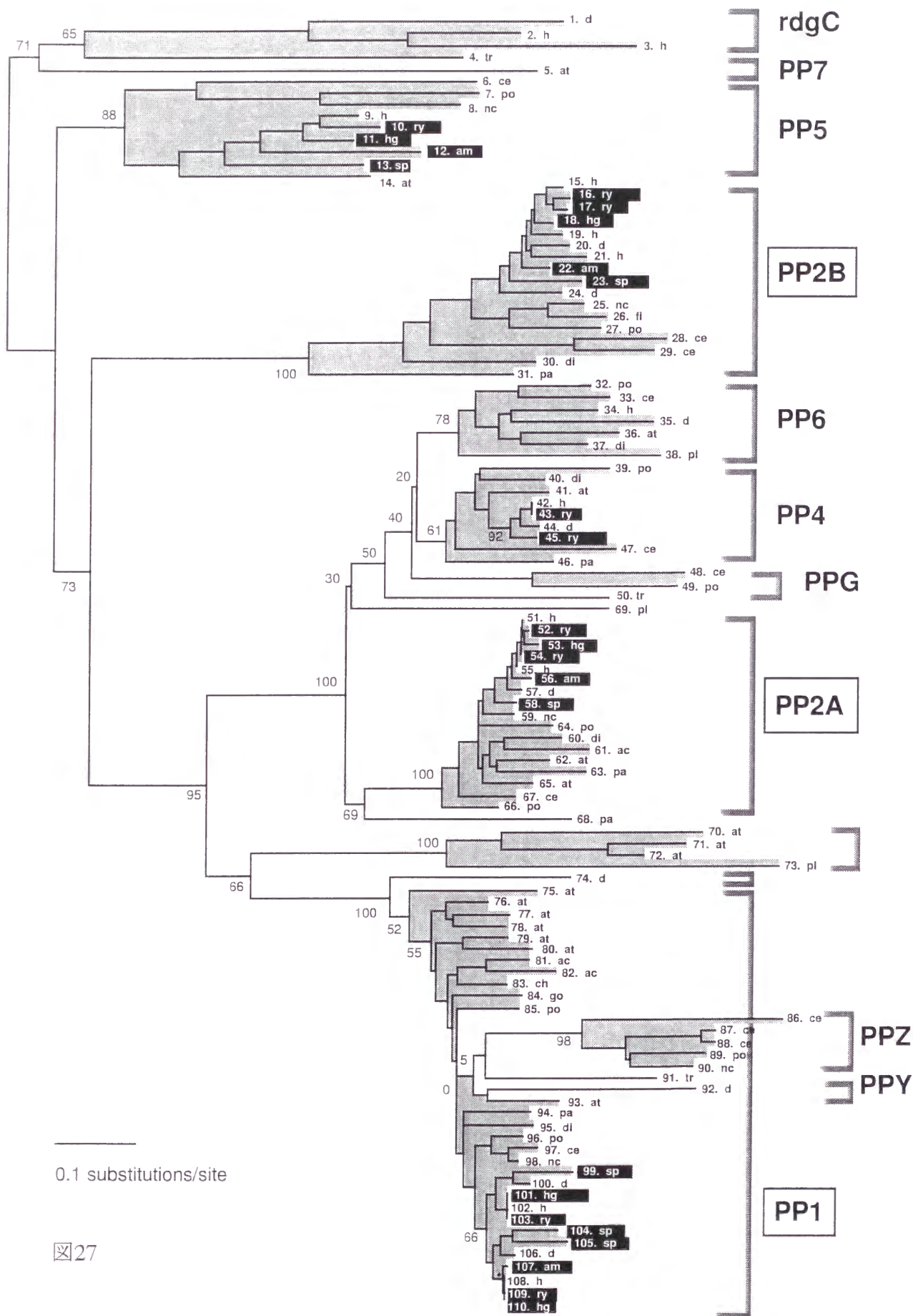


図27 PP遺伝子族の分子系統樹

PPドメインの約3分の2に相当するアライメント領域（168アミノ酸座位）に基づき、近隣結合法により推定した分子系統樹を示した。アライメントは付録5-9に準ずる。単離されたカワカイメン・ナメクジウオ・ヌタウナギ・エイのPP遺伝子は白抜き文字で示した。各系統の記号は以下を表す。h, human; ry, ray; hg, hagfish; am, amphioxus; d, *Drosophila*; s, sponge; ce, *Saccharomyces cerevisiae*; po, *Schizosaccharomyces pombe*; fi, *Filobasidiella*; nc, *Neurospora*; ac, *Acetabularia*; ch, *Chlamydomonas*; at, *Arabidopsis*; di, *Dictyostelium*; go, *Gonyaulax*; pa, *Paramecium*; pl, *Plasmodium*; tr, *Trypanosoma*。系統樹推定に用いた遺伝子の名前及びGenBankのアクセッションナンバーは付録3-2に示した。同じサブタイプに属する遺伝子群を網掛けで示し、サブタイプ名を右側に記した（サブタイプ名のないものは記していない）。アイソフォームの作り出された時期について詳細な解析を行うサブタイプを四角で囲った。尚ギアルディアの配列を含めると、さらにアライメント領域が減少し系統樹の信頼性が低下するため、ギアルディアの配列は除いて系統樹推定を行った。ノード付近に示した数字はブートストラップ確率を示す。

PP1サブタイプが単系統群を形成しないのは、系統樹推定に用いたアライメント領域が減少したことによる統計的なゆらぎだと思われる。また、菌類の遺伝子群が単系統群を形成しないサブタイプや、菌類・細胞性粘菌・植物の分岐の順序が正しい分岐順序を反映していないサブタイプが見受けられる。このことは、この遺伝子族が保存的であるために、進化速度の違いによる影響を受けやすかったり、接近した分岐である菌類・細胞性粘菌・植物の分岐を正確に推定出来なかった結果であると考えられる。生物種の系統関係を反映していないこれら菌類・細胞性粘菌・植物の遺伝子のホモログだと思われる動物の遺伝子が存在しないことから考えても、これらは遺伝子重複ではなく生物種の分岐が誤って推定されたものとするのが自然だと思われる。

PP遺伝子族のアイソフォームが作り出された時期の推定

A. PP2Aサブタイプ

脊椎動物には、PP2A α とPP2A β の2種類のアイソフォームが存在する。本論文では新たに、カワカイメンから1種類、ナメクジウオから1種類、ヌタウナギから1種類、エイから2種類の、PP2Aサブタイプに属する遺伝子を単離した。これらの遺伝子を併せて、網羅的探索による最尤法を用いて、PP2Aサブタイプの分子系統樹を推定した。アウトグループには、PP2Aサブタイプに属するカイメンのsPP2A遺伝子を用いた。その結果、図28の分子系統樹が最尤系統樹となった。

エイのryPP2Abは、ヒト及びアフリカツメガエルのPP2A α アイソフォームと、分子系統樹推定に用いた計算領域においてアミノ酸配列が完全に一致した。したがって、このryPP2AbはPP2A α アイソフォームであると考えられる。また分子系統樹から、エイのryP2AaとヌタウナギのhgPP2Aは、新しいアイソフォーム（ry/hgPP2Aとする）だと考えられた。ナメクジウオのamPP2Aは、3種のアイソフォームとオーソログスな関係にある遺伝子である。

3種類のアイソフォームを作り出した遺伝子重複2回（菱形で示す）は、いずれもナメクジウオと脊椎動物の分岐（二重丸）後、ヌタウナギと有顎類の分岐（黒丸）前に起きている。

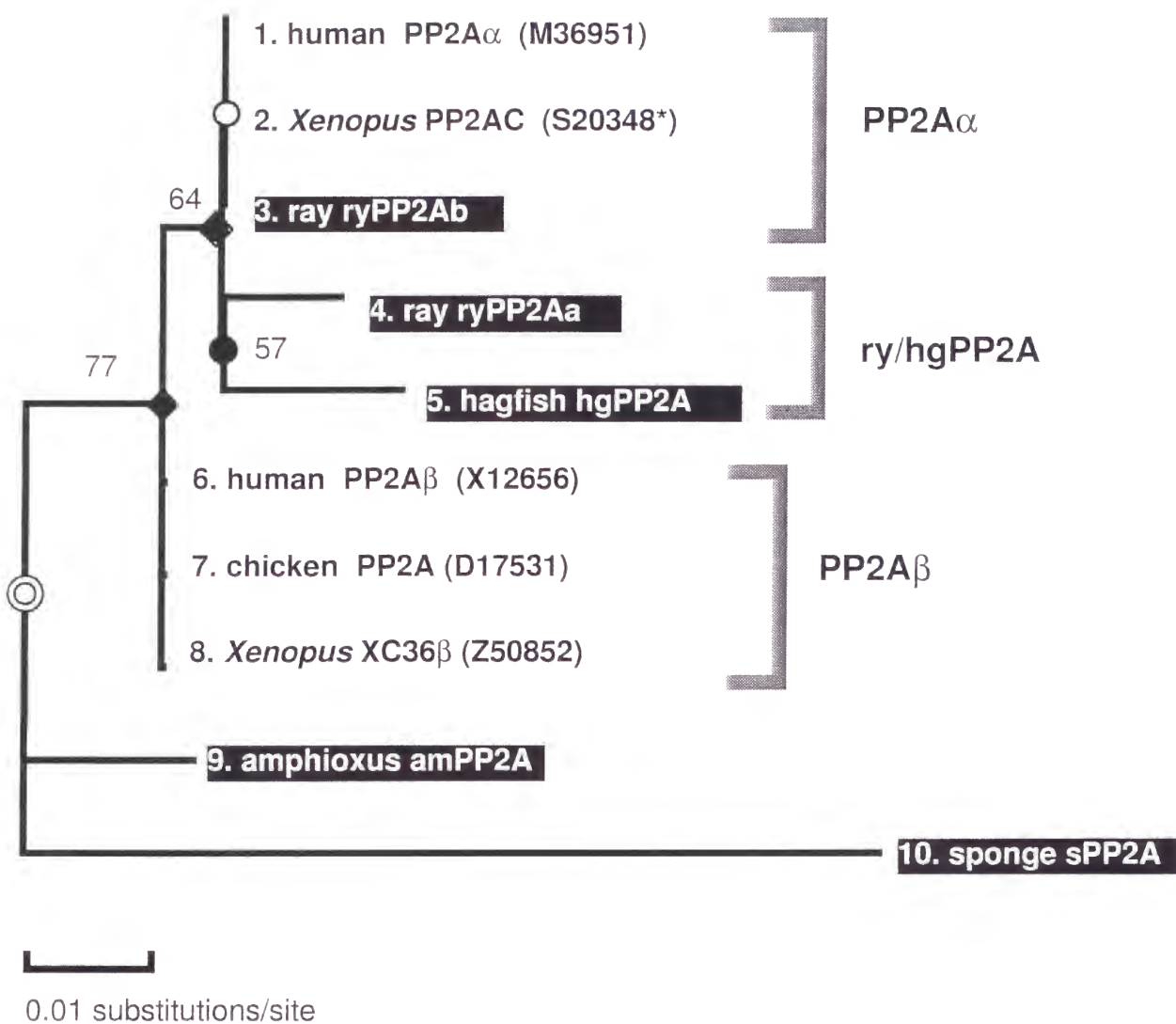


図28 PP2Aサブタイプの分子系統樹

PPドメインの約3分の2に相当するアライメント領域（219アミノ酸座位）に基づき、網羅的探索による最尤法で推定された最尤系統樹を示した。尤度の等しい3つの系統樹が最尤系統樹となったが、このうち最もブートストラップ確率の高い（45.9%。残りの二つはそれぞれ5%と0.1%。）樹形を示した。アライメントは付録5-10に示した。アウトグループには、PP2Aサブタイプに属するカイメンのsPP2Aを用いた。本論文で決定した配列は白抜き文字で示した。配列には生物種名、遺伝子名、GenBankのアクセッションナンバー（*のついたものはPIRのアクセッションナンバー）を記載し、アイソフォームの名前を右側に記した。ノード付近に示した数字はブートストラップ確率である。この最尤系統樹と尤度の差が1標準誤差範囲内である樹形が他に29通り（105通り中）存在した。

二重丸は頭索動物と脊椎動物の分岐を、黒丸は無顎類と有顎類の分岐を、白丸は軟骨魚類とその他の有顎類の分岐を、菱形は遺伝子重複を表す。

B. PP2Bサブタイプ

脊椎動物には、PP2B α 、PP2B β 、PP2B γ の3種類のアイソフォームが存在する。本論文では新たに、ナメクジウオから1種類、ヌタウナギから1種類、エイから2種類の、PP2Bサブタイプに属する遺伝子を単離した。これらの遺伝子を合わせて、網羅的探索による最尤法を用いて、PP2Bサブタイプの分子系統樹を推定した。アウトグループには、PP2Bサブタイプに属するショウジョウバエの遺伝子Pp2B-14DとCanA1を用いた。その結果、図29の分子系統樹が最尤系統樹となった。

図29の分子系統樹から、エイのryPP2BaはPP2B α アイソフォームに、ヌタウナギのhgPP2BはPP2B β アイソフォームに分類された。エイのryPP2Bbは新しいアイソフォームだと考えられる。ナメクジウオのamPP2Bは、4種のアイソフォームとオーソログスな関係にある遺伝子である。

この4種類のアイソフォームを作り出した遺伝子重複3回（菱形で示す）は、いずれもナメクジウオと脊椎動物の分岐（二重丸）後に起きていた。そのうち1回は、ヌタウナギと有顎類の分岐（黒丸）前に起きていた。残りの2回の遺伝子重複が起きた時期は、ヌタウナギの分岐前か後かは不明であるが、少なくともエイとその他の有顎類の分岐（白丸）前である。

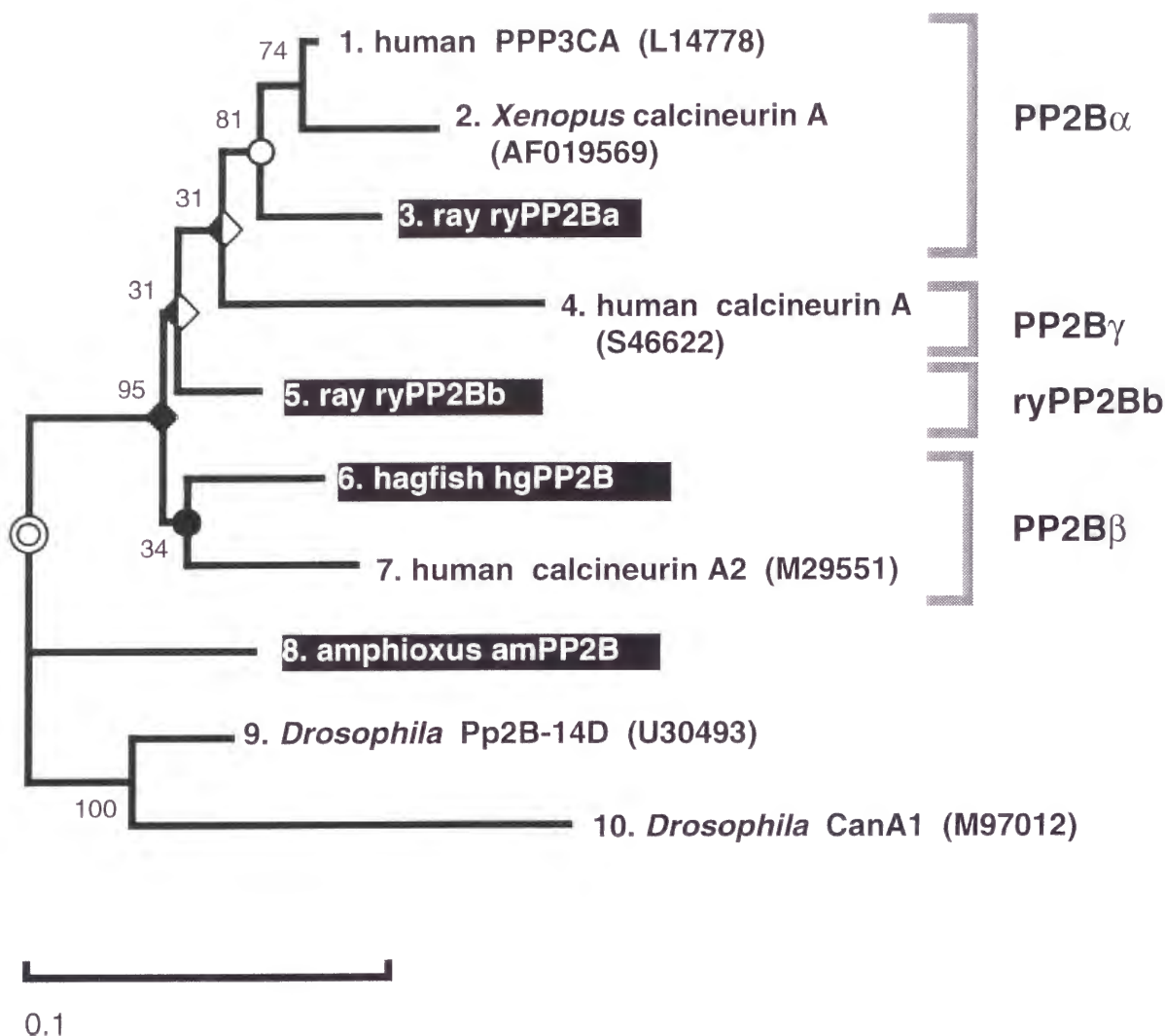


図29 PP2Bサブタイプの分子系統樹

PPドメインの約3分の2とそのC末側に存在するカルモジュリン結合部位を含むアライメント領域 (340 アミノ酸座位) に基づき、網羅的探索による最尤法で推定した最尤系統樹を示した。アライメントは付録5-11に示した。アウトグループにはPP2Bサブタイプに属するショウジョウバエのPp2B-14D及びCanA1遺伝子を用いた。本論文ではカワカイメンのPP2Bサブタイプに属する遺伝子も単離しているが、ショウジョウバエの遺伝子の方がより近縁であり、信頼性の高い系統樹推定が可能だと考え、結果が変わらないことを確認した上で、ショウジョウバエの遺伝子をアウトグループに用いた。本論文で決定した配列は白抜き文字で示した。配列には生物種名、遺伝子名、GenBankのアクセッションナンバーを記載した。ノード付近に示した数字はブートストラップ確率である。この最尤系統樹と尤度の差が1標準誤差範囲内である樹形が他に110通り (135135通り中) 存在した。

二重丸は頭索動物と脊椎動物の分岐を、黒丸は無顎類と有顎類の分岐を、白丸は軟骨魚類とその他の有顎類の分岐を表す。黒い菱形は、頭索動物と脊椎動物の分岐後、無顎類と有顎類の分岐前に起きた遺伝子重複を、白黒の菱形は、頭索動物と脊椎動物の分岐後、少なくとも軟骨魚類とその他の有顎類の分岐前に起きた遺伝子重複を表す。

C. PP1サブタイプ

脊椎動物にはPP1サブタイプに属する遺伝子が3種類(PP1 α 、PP1 β 、PP1 γ)存在する。本論文では新たに、カワカイメンから3種類、ナメクジウオから1種類、ヌタウナギから2種類、エイから2種類の、PP1サブタイプに属する遺伝子が単離された。これらの遺伝子を併せて、近隣結合法によりPP1サブタイプの分子系統樹を推定した(図30)。その結果、アイソフォームが作り出された時期が2期あることがわかった。

脊椎動物のPP1サブタイプに属するアイソフォーム3種類のうち、PP1 β とPP1 α /PP1 γ アイソフォームが分岐したのは、海綿動物とその他の動物の分岐(黒丸)前であった(黒い菱形)。発見的探索による最尤法で得られた最尤系統樹でも、このことが示された。

PP1 β と、PP1 α /PP1 γ アイソフォームの分岐後、PP1 α とPP1 γ が分岐した。その時期を詳細に推定する目的で、PP1 α 及びPP1 γ アイソフォームに属する遺伝子群のみを選び出し、網羅的探索による最尤法を用い分子系統樹を推定しなおした。アウトグループには、ホヤの遺伝子HcPP1を用いた。その結果、図31の分子系統樹が最尤系統樹となった。

エイのryPP1aはヒトのPPP1CCと、分子系統樹推定に用いた計算領域においてアミノ酸配列が完全に一致した。したがって、ryPP1aはPP1 γ アイソフォームに属すると考えられる。また、ヌタウナギのhgPP1a及びアフリカツメガエルのPP1 γ もPP1 γ アイソフォームだと考えられる。PP1 γ アイソフォームに属する4種の生物の分岐が、系統関係を反映していないのは、PP1 γ アイソフォームが保存的な分子であることに起因すると思われる。ナメクジウオのamPP1は、PP1 α アイソフォーム及びPP1 γ アイソフォームとオーソログな関係にある遺伝子である。

この2種類のアイソフォームを作り出した遺伝子重複(菱形で示す)は、ナメクジウオと脊椎動物の分岐(二重丸)後、ヌタウナギと有顎類の分岐(黒丸)前に起きている。

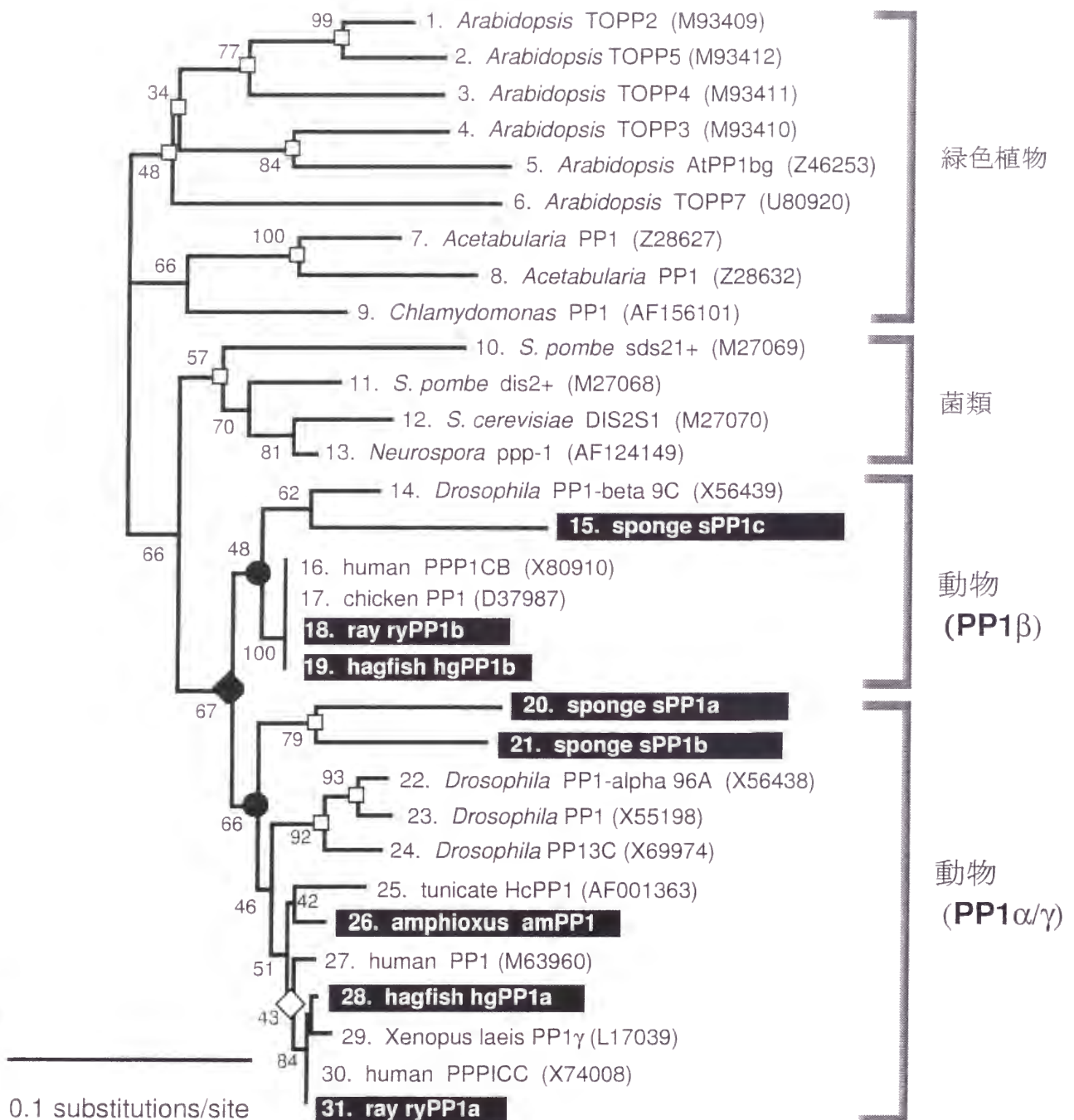


図30 PP1サブタイプの分子系統樹 (PP1α, PP1β, PP1γアイソフォーム)

PPドメインの約3分の2に相当するアライメント領域 (205アミノ酸座位) に基づき、近隣結合法により推定した分子系統樹を示した。アライメントは付録5-12に示した。本論文で決定した配列は白抜き文字で示した。配列には生物種名、遺伝子名、GenBankのアクセッションナンバーを記載した。ノード付近に示した数字はブートストラップ確率である。PP1βの無脊椎動物及びPP1αの原索動物の分岐が、生物種の系統関係を反映しないのは、この分子が保存的であることに起因する統計的ゆらぎだと思われる。

黒丸は海綿動物と真正後生動物の分岐を表す。黒の菱形は菌類と動物の分岐後海綿動物と真正後生動物動物の分岐前に起きた遺伝子重複を、白の菱形は海綿動物と真正後生動物の分岐後に起きた遺伝子重複を表す。四角は脊椎動物に至る系統以外で起きた遺伝子重複を表す。

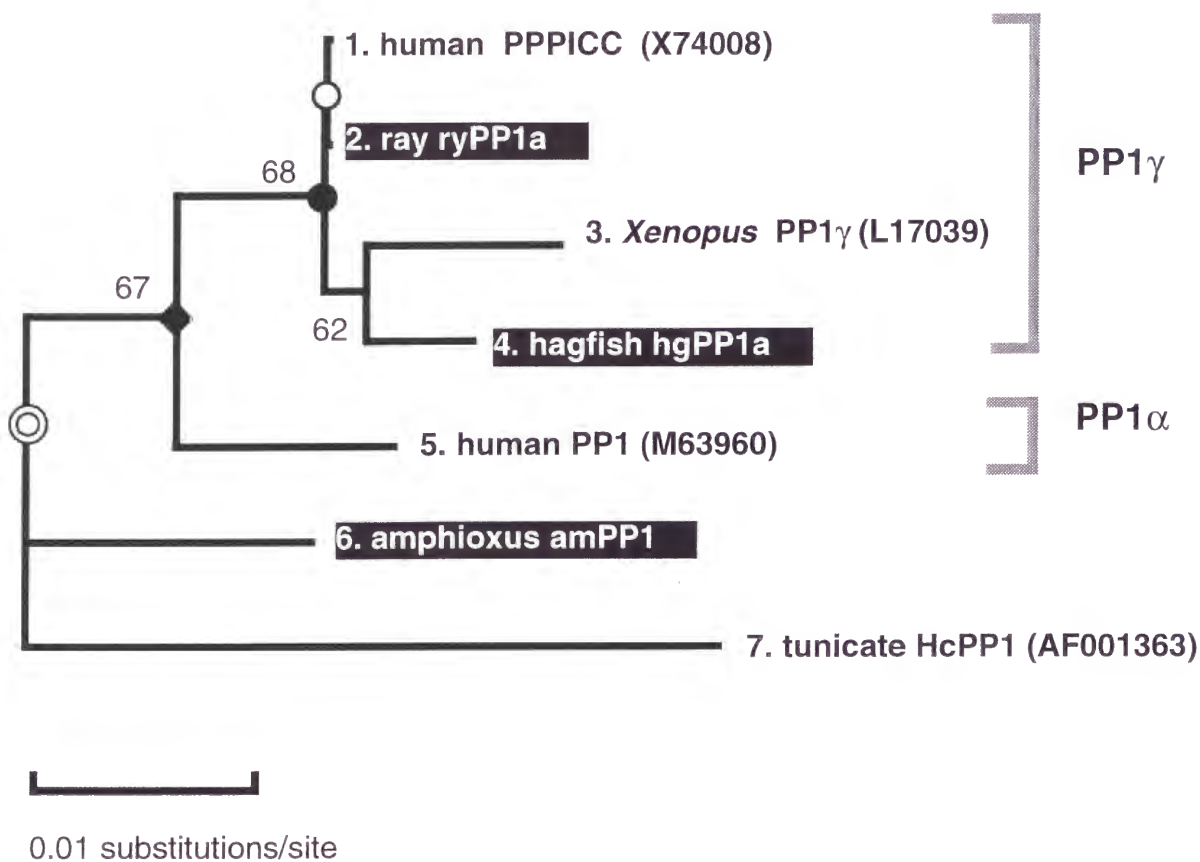


図31 PP1サブタイプの分子系統樹 (PP1 α , PP1 γ アイソフォーム)

PPドメインの約3分の2に相当するアライメント領域 (206アミノ酸座位) に基づき、網羅的探索による最尤法で推定された最尤系統樹を示した。アライメントは付録5-12に準ずる。アウトグループには、PP1サブタイプに属するホヤのHcPP1遺伝子を用いた。本論文で決定した配列は白抜き文字で示した。配列には生物種名、遺伝子名、GenBankのアクセッションナンバーを記載し、アイソフォームの名前を右側に記した。ノード付近に示した数字はブートストラップ確率である。この最尤系統樹と尤度の差が1標準誤差範囲内である樹形が他に20通り (105通り中) 存在した。

二重丸は頭索動物と脊椎動物の分岐を、黒丸は無顎類と有顎類の分岐を、白丸は軟骨魚類とその他の有顎類の分岐を、菱形は遺伝子重複を表す。

PP遺伝子族のアイソフォームを作り出す遺伝子重複が起きた時期の再検討

以上3つのサブタイプにおいて、アイソフォームを作り出した遺伝子重複が起きた時期は、PP1サブタイプで見られた一例を除き、ナメクジウオと脊椎動物の分岐後、エイとその他の有顎類の分岐前であった。これらの遺伝子重複のうちヌタウナギと有顎類の分岐との関係については、ヌタウナギと有顎類の分岐前に起きたものが多かった。

3つのサブタイプの最尤系統樹推定の際には、図28、図29、図31で示した最尤系統樹の他に、一標準誤差範囲内に棄却できない複数の樹形が存在した。そこで、PTP遺伝子族で行った解析と同様の解析を行った。すなわち、一標準誤差範囲内にある樹形全てについて、アイソフォームを作り出した遺伝子重複が起きた時期及びその回数を推定し、各樹形のブートストラップ確率で重みづけをして、遺伝子重複が起きた時期について再検討した（材料と方法の章参照）。その結果、無顎類と有顎類の分岐前（III-b期）に起きた遺伝子重複は計4.0回、これに対して無顎類と有顎類の分岐後（III-c期）に起きた遺伝子重複は計1.3回となった。また、図28、図29、図31で得られた結果と同様、頭索動物と脊椎動物の分岐前（III-a期）、あるいは軟骨魚類とその他の脊椎動物の分岐後（IV期）には、遺伝子重複はほとんど見られないことが確かめられた（表15）。

表15 PP遺伝子のアイソフォームを作り出す遺伝子重複が起きた時期及びその回数

サブタイプ	III-a期	III-b期	III-c期	IV期
PP2A	0.26（回）	1.1	0.50	0.0
PP2B	0.0	1.3	0.74	0.0
PP1	0.22	1.6	0.01	0.0

III-a期は、ナメクジウオと脊椎動物の分岐前の時期、III-b期は、ナメクジウオと脊椎動物の分岐後ヌタウナギと有顎類の分岐前の時期、III-c期はヌタウナギと有顎類の分岐後エイとその他の有顎類の分岐前の時期、IV期はエイとその他の有顎類の分岐後現在までを指す。

考察

結果のまとめ

RT-PCRによりカワカイメン・ナメクジウオ・ヌタウナギ・エイから計24種類のPP遺伝子が単離された。

単離した遺伝子を含めて分子進化学的解析を行った結果、PP1, PP2A, PP2Bの3つのサブタイプのアイソフォームを作り出した遺伝子重複が起きた時期は、PP1サブタイプの一例を除いて、ナメクジウオと脊椎動物の分岐後、エイのその他の有顎類の分岐前の、ヌタウナギの分岐の周辺であることがわかった。PP1サブタイプのPP1 β とPP1 α /PP1 γ アイソフォームを作り出した遺伝子重複は、菌類と動物の分岐後、海綿動物と真正後生動物の分岐前であった（図32）。

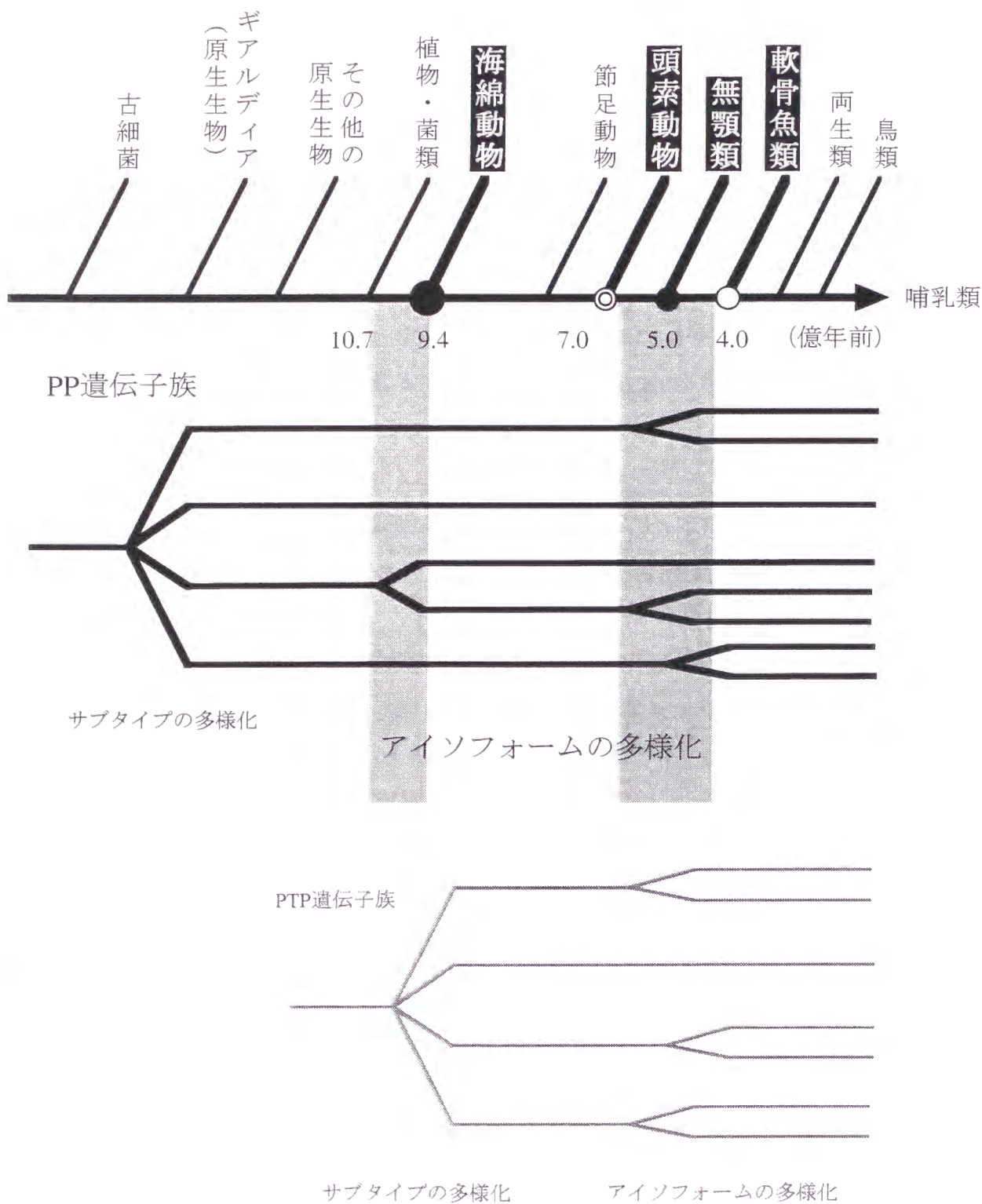


図32 PP遺伝子族のアイソフォームの多様化のまとめ

PP遺伝子族のアイソフォームの多様化の時期を模式的に示した。本論文の結果から明らかとなった、アイソフォームの多様化が見られた時期を網掛けで示した。枝の長さは概ね時間を反映しているが正確ではない。数字は分岐年代を示す(Dickerson, 1971; Dayhoff, 1978; Nikoh et al., 1997)。比較のため、下にPTP遺伝子族の多様化が見られた時期（前節1-1及び1-2参照）を示した。

PP遺伝子族とPTP遺伝子族の結果の比較

PP遺伝子族でアイソフォームの多様化がみられた時期は、PTP遺伝子族でアイソフォームの多様化のみられた時期（脊椎動物進化の初期）と一致していた。また、PP1サブタイプについては、PTP遺伝子族でサブタイプの多様化のみられた時期（多細胞動物進化の初期）にも、アイソフォームを作り出す遺伝子重複が起きていた。この遺伝子重複では、PP1 β とPP1 α /PP1 γ が作り出された。現在のところ、これらのアイソフォーム間の差異は明確でないが、PTP遺伝子族の結果を考慮すると、多細胞化に関係するような機能の差異が存在するのかもしれない。いずれにしても、PP遺伝子族のアイソフォームの多様化は、限定した時期に起きており、その時期はPTP遺伝子族で多様化がみられた時期と一致していた。

PTP遺伝子族において、サブタイプやアイソフォームが作り出された時期は、それ以外の時期と比べて分子進化速度が速かった。PP遺伝子族にも同様の傾向がみられるだろうか。

PTP遺伝子族でサブタイプの多様化のみられた多細胞動物進化の初期には、上述したようにPP1サブタイプでアイソフォームが作り出されていた。PP1サブタイプにおいて、この遺伝子重複が起きた時期は、それ以降と比べて4.3倍分子進化速度が速かった。したがってこの時期には遺伝子重複だけでなく、アミノ酸置換も多く起きていたことがわかる。

PTP遺伝子族でアイソフォームの多様化のみられた脊椎動物進化の初期には、PP遺伝子族でも分子進化速度が高かった。あるいは、脊椎動物の系統で分子進化速度が急激に低くなっているといえた。例えば、PP2Aサブタイプにおいて、ヒト・アフリカツメガエル・エイのPP2A α アイソフォームに属する遺伝子は、分子系統樹推定に用いた計算領域において、アミノ酸配列が完全に一致した。この他にもPP2A β アイソフォーム、PP1 β アイソフォーム、PP1 γ アイソフォームにおいて、同様に脊椎動物の間でアミノ酸配列が完全に一致していた。このような脊椎動物の系統における分子進化速度の低下は、他の遺伝子族についても報告されている(Sidow, 1992; Hoshiyama et al., 1998; Hoshiyama et al., manuscript in preparation)。このことについては総合議論の章で考察することにする。

PP遺伝子族のアイソフォームの多様化と生物進化との関連性

PP遺伝子族でアイソフォームの多様化がみられた時期は、多細胞動物進化の初期及び脊椎動物進化の初期の限定した時期であった。この時期は動物の初期進化において多細胞性が獲得された時期、脊椎動物の初期進化における器官や組織の進化が起きた時期に対応する。

一方、それ以外の時期には遺伝子重複はほとんど起きていないことが明らかとなった。したがって、PP遺伝子族のアイソフォームの多様化と、例えば多様な動物門が分岐したカンブリア爆発とは、関連の無いことが示唆される。

動物以外の系統で起きたアイソフォームの多様化

本論文では、動物の系統で起きたアイソフォームの多様化にのみに注目した。しかし、PP遺伝子族の分子系統樹は、植物や菌類の系統でもそれぞれ独自にアイソフォームの多様化が起きたことを示している。例えばシロイヌナズナには、PP1サブタイプに属する複数種類の遺伝子（例えばTOPP2~5など）が存在する。図30の分子系統樹から、これらのシロイヌナズナのアイソフォームが遺伝子重複により作り出された時期は、少なくともクラミドモナスやカサノリといった単細胞の緑藻類の分岐後であることがわかる。これらのアイソフォームの多様化が起きた詳しい時期はいつか、そしてそれは植物の形態の進化とどう関連しているのかを探ることは、今後の課題である。

総合議論

1. 結論のまとめ

プロテインホスファターゼ遺伝子族の多様化に関する分子進化学的解析から、以下の2点が明らかとなった（図33）。

1) 脊椎動物に至る進化の過程において、遺伝子重複は断続的に起きていた。すなわち、遺伝子重複の見られた時期は、真核生物進化の初期（A'期）、多細胞動物進化の初期（B'期、10.7億年前～9.4億年前、Nikoh et al., 1997）、脊椎動物進化の初期（C''期、7億年前～4億年前、Dayhoff, 1978; Nikoh et al., 1997）、の3期に限定していた。

2) A'期、B'期、C''期のそれぞれの時期には、異なる特徴をもった遺伝子が作り出されていた。すなわち、A'期に起きた遺伝子重複では、PPの機能の異なるサブタイプが作り出されていた。B'期に起きた遺伝子重複では、主にPTPの機能の異なるサブタイプが作り出されていた。C''期に起きた遺伝子重複では、PP及びPTPともアイソフォームが作り出されていた。

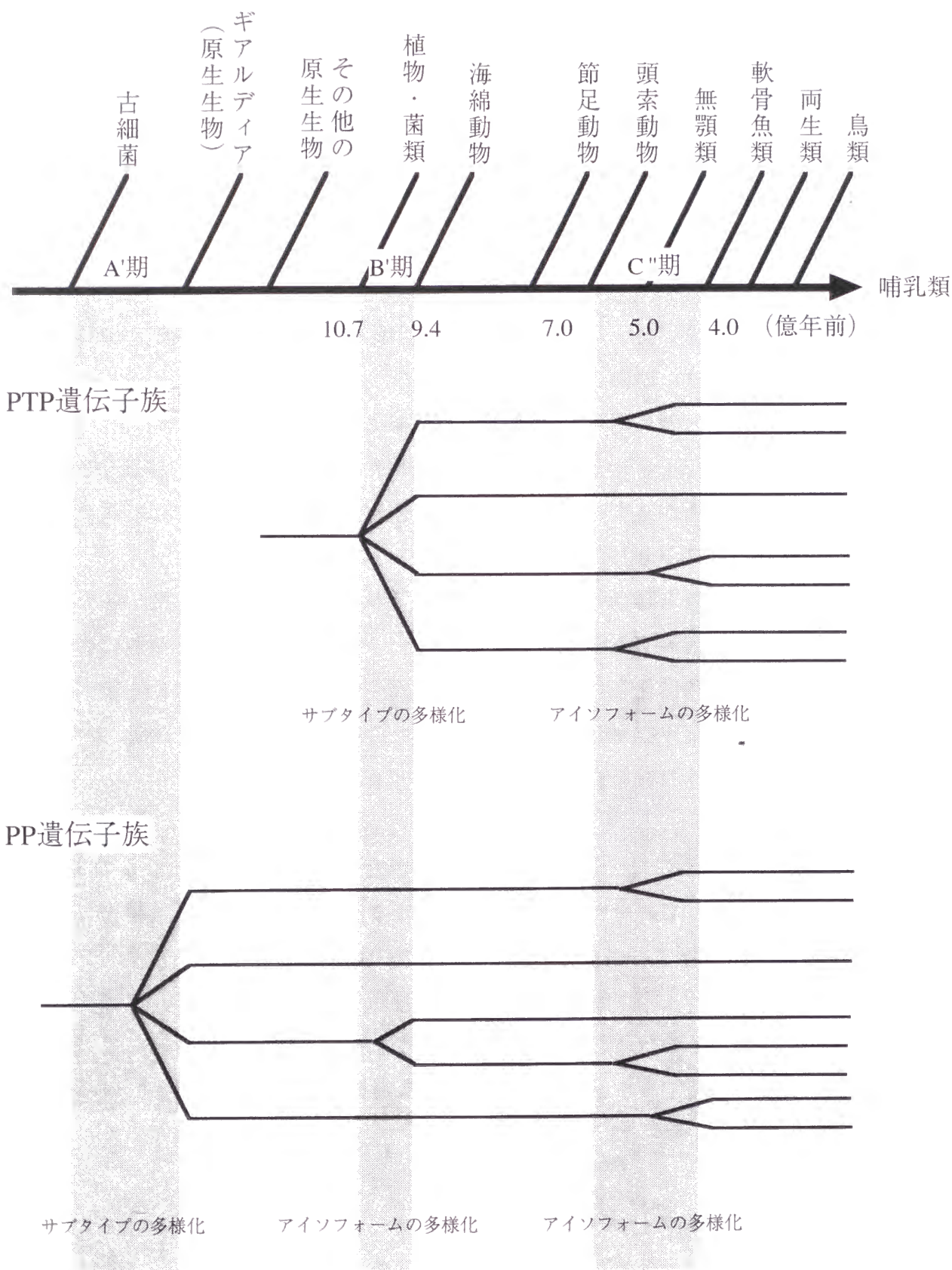


図33 結論のまとめ

本論文の結果から明らかとなった、プロテインホスファターゼ遺伝子族の多様化を模式的に示した。枝の長さは概ね時間を反映しているが正確ではない。数字は分岐年代を示す(Dickerson, 1971; Dayhoff, 1978; Nikoh et al., 1997)。

以上の結果から、遺伝子重複は生物の進化にともなって、次々に起きているのではなく、ある時期に限定して起きていることが明らかとなった。したがって、遺伝子重複と生物の形態進化とは、1)関連性の認められる面と、2)関連性の認められない面のあることが考察された。すなわち、断続的な遺伝子重複がみられた時期に相当する生物の進化は、遺伝子重複と何らかの関連性のある可能性が示唆された。一方、遺伝子重複がほとんどみられない時期に相当する生物の進化は、遺伝子重複と関連性がないと考えられた。(図34)

1) A'期は、真核生物が核を獲得し、細胞骨格や細胞内小器官などの細胞内構造が進化した時期に相当する。またB'期は、動物が多細胞性を獲得した時期である。そしてC''期は、脊椎動物の神経系・消化器系・循環器系などの独立した組織や器官が進化した時期に相当する。このように断続的な遺伝子重複がみられた時期は、生物が新しい構造や体制を獲得した時期に対応している。したがって、このような生物進化がみられた時期には、遺伝子重複が重要な役割を果たしたのかもしれない。

もちろん、遺伝子重複が起きた時期と形態上の進化が見られた時期が一致しているからといって、必ずしもその間に因果関係があるとは限らない。今後、例えば海綿動物において、多細胞動物進化の初期に多様化した遺伝子が、多細胞性に寄与しているのかを調べることなどは重要である。そのような研究により、断続的な遺伝子重複とA'期、B'期、C''期にみられた生物の進化の関連性の有無を検証することが出来るだろう。

2) A'期・B'期・C''期の3期以外には、遺伝子重複はほとんど起きていなかった。このような時期にも生物の進化は滞っていたわけではない。多様な原生生物が出現し、植物・菌類が出現し、最も劇的な出来事として多様な多細胞動物が出現したカンブリア爆発も起き、陸上脊椎動物が進化した。このような進化は、その時期のずれから考えて、遺伝子重複とは関連性が無いように思われた。したがってこのような多様な生物種の分岐などの生物進化は、新しい遺伝子を創造することなく達成されたと考えられる。このことについては後で詳しく考察することにする。

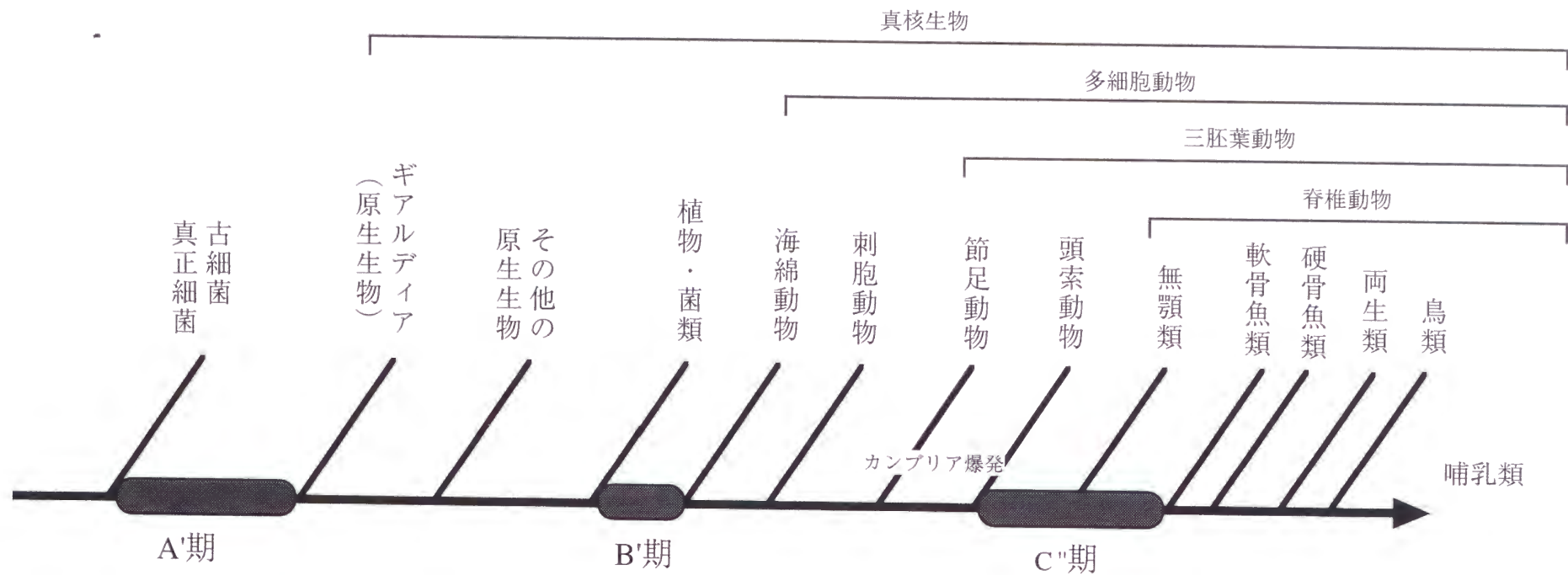


図34 遺伝子重複と生物の進化

本論文の結果をもとに、遺伝子重複が起きた時期と生物の進化の対応を模式的に示した。枝の長さは概ね時間を反映しているが正確ではない。正確な分岐年代は材料と方法の章参照。プロテインホスファターゼ遺伝子族で断続的な多様化がみられた時期を網掛けで示した。

2 遺伝子族の多様化に関する今後の課題

岩部らの解析(Iwabe et al., 1996)から他のいくつかの遺伝子族についても、遺伝子重複が起きた時期が少なくとも、A期、B期、C期の三つの時期に分類されることが示されている(序論4参照)。これらの遺伝子族の多様化も、プロテインホスファターゼ遺伝子族で多様化のみられた時期すなわち、A'期、B'期、C''期に限定しているのか、今後さらに多くのデータを蓄積し検証していく必要があるだろう。

また岩部らの解析した遺伝子族以外の遺伝子族にも、同様の断続的多様化がみられるのだろうか。すなわちこの現象は全ての遺伝子族に共通したものなのだろうか。例えば、動物の形態進化との関連を考える上で、形態形成に深く関わる遺伝子族や転写因子の遺伝子族に、同様の傾向がみられるか否かは興味深い問題である。このような遺伝子族の中にはこれまでの結果と異なり、例えばカンブリア爆発の起きた時期に多様化のみられた遺伝子族があるのかもしれない。ただし星山らの解析から、形態形成に関わる転写因子であるPax遺伝子族に関しては、PTP遺伝子族と同様の傾向がみられることが示されている(Hoshiyama et al., 1998)。したがって、形態形成に深く関わる遺伝子族についても、カンブリア爆発の起きた時期には多様化は見られなかったのかもしれない。

いくつかの遺伝子族については、A期、B期、C期の三期以外にも多様化がみられることが報告されている。例えば、ATPase associated with diverse cellular activities (AAA)遺伝子族では、真正細菌との分岐後、古細菌と真核生物の分岐前の時期に多様化がみられる(Swaffield & Purugganan, 1997)。またインターフェロンや成長ホルモンの遺伝子族では比較的最近、すなわち鳥類の分岐後哺乳類の進化の過程で遺伝子重複がみられるようである(Roberts et al., 1998; Forsyth, 1994 for review)。これらの遺伝子族の多様化と生物進化との関連も興味深い問題である。

また本論文では、脊椎動物に至る系統で起きた遺伝子族の多様化に限定して解析を行ったが、それ以外の系統、すなわち植物や菌類あるいは無脊椎動物の系統でも、断続的な遺伝子重複が見られるのだろうか。PP遺伝子族では、植物・菌類の各系統でも遺伝子重複がみられた。またPTP遺伝子族のあるサブタイプでは、カワカイメンの系統において非常に多くの遺伝子重複がみられた。これらの遺伝子重複の起きた時期、及び形態進化との関連性の検証は今後の課題である。

現在様々な生物において、全ゲノムの塩基配列を解読するという試みが盛んに行われている。これらのデータを用いた解析から、遺伝子族の多様化が起きた時期の全貌を明らかにすることが出来るだろう。またそのためには、本論文で取り扱ったような進化的に重要な生き物(例えば海綿動物、無顎類など)の全ゲノムの塩基配列の決定が必須である。

3. 新しい遺伝子を創造することなく多様な生物が進化するしくみ

本論文の結果から、遺伝子重複と生物進化の間には関連性の認められない面もあることが考察された。例えば、多様な多細胞動物が進化したカンブリア爆発の起きた時期には、遺伝子族のサブタイプの多様化もアイソフォームの多様化も見られなかった。それでは、新しい遺伝子を創造することなく、多様な生物が進化するしくみにはどのようなものが考えられるだろうか。

序論で形態進化と分子進化の間に存在する溝について述べた（序論2参照）。最近、この溝を埋めるヒントとなる結果がRutherfordらにより示された(Rutherford & Lindquist, 1998)。Rutherfordらは、熱ショックタンパク質の一種であるHSP90に、形態異常の原因となる多種類の突然変異を遮へいする効果があることを示した。HSP90の機能下では、突然変異により引き起こされる異常が隠されるため、突然変異は「中立」な変異として集団に広まることが出来る。このような突然変異が蓄積しある閾値を越えた時、あるいは環境の変化が起きた時、HSP90の機能で異常を覆いきれなくなり、複数の突然変異によりもたらされる形態上の変化が露出するという。Rutherfordらは、このような機構が革新的な形態進化の説明になるのかもしれないと論じている。このRutherfordらの結果は、突然変異の役割を再認識させるものだともいえる。

それでは新しい遺伝子を創造することなく突然変異により、如何にして機能の変化あるいは新機能の獲得が達成されたのだろうか。遺伝子をコードする領域に起きた突然変異は、個体に悪影響の無い範囲で(HSP90の効果も含め) 分子の生化学的性質の変化を許容し、分子の新しい機能の獲得をもたらしかもしれない。また、相互作用する相手分子が変化することで生化学的反応経路が変化し、これにより新しい機能が獲得されることもあるだろう。例えば、昆虫の外肢形成で働くHOX遺伝子の下流分子が昆虫の進化の過程で変化していること(Palopoli & Patel, 1998)、線虫の陰門形成で働く同じ遺伝子が種間で異なる反応経路をもつこと(Eizinger et al., 1999 for review)、などは生化学的反応経路の変化の具体例である。一方、調節領域に起きた突然変異は、分子をいつ、どの組織で発現させるかという発現の時間的・空間的多様性を生み出し、これにより新しい機能獲得の機会が得られるだろう。実際、発生に関わる *pax6* (Strachan and Read, 1994; Mansouri et al., 1996) や *engrailed* (Patel et al., 1989) などの分子は、発生の過程で繰り返し多目的に使用されることが知られている。また、全く同じ分子を違う目的で使うようになった例も知られている。レンズクリスタリンが酵素と兼用されていること(Wistow et al., 1990)、伸長因子であるEF1 α が細胞骨格の調節も行うこと(Condeelis, 1995 for review)、などがその例である。星山ら

は、以上に述べたような遺伝子の機能変化あるいは新機能獲得の機構の存在を示唆する間接的な証拠を得ている。すなわち、いくつかの遺伝子族において脊椎動物の系統で分子進化速度の低下が見られ(Sidow, 1992; Hoshiyama et al., 1998)、このことが生化学的反応経路の複雑化あるいは時間的・空間的発現の複雑化に伴う機能的制約の強化に起因するのではないかと考察している(Hoshiyama et al., 1998; Hoshiyama et al., manuscript in preparation)。

遺伝子の機能変化あるいは新機能の獲得は、gene recruitment (あるいはco-optin) とも呼ばれ、多様な生物が進化した原動力の一つと考えられる(e.g. Hoshiyama et al., 1998; Ganfornina & Sanchez, 1999)。このgene recruitmentという考え方は、今後生物の進化を分子レベルで理解する上で重要な鍵であろう。

4. 展望

本論文の結果から、遺伝子族の多様化と生物進化の関連性を検証するためには、本論文で取り扱ったような生物（海綿動物、無顎類など）における重複遺伝子の機能の解明が重要であると考えられた。また、遺伝子族の多様化と生物進化との間には関連性のない面があることも明らかになり、新しい遺伝子を創造することなく多様な生物が進化したしくみを理解するためには、多様な生物種間で遺伝子がどう使われているかに目を向けることが重要であると考えられた。いずれにせよ、生物進化の分子基盤を解明するためには、異なる生物種間での遺伝子の機能の比較が重要である。その際には、個々の遺伝子の機能に注目するだけでなく、その遺伝子が機能する生化学的反応経路、発現時期及び場所といった、遺伝子の機能するシステム全体を包括的に捉える視点が重要だと考えられる。今後、生物種間でこのシステムを比較することができれば、生物進化の分子基盤を理解するための有益な情報が得られるだろう。

最近この問題にアプローチできるような技術が現れた。2次元電気泳動(プロテオーム、e.g. James, 1997 for review)やDNAマイクロアレイ(DNAチップ、e.g. Lander, 1999 for review)といった技術は、包括的に遺伝子発現を捉えることを可能にした。特にDNAマイクロアレイは、ヒトの多型の検出(Wang et al., 1998)や、がん細胞における異常遺伝子の探査(Perou et al., 1999)、さらには環境変化に伴う遺伝子発現変化のモニタリング(DeRisi et al., 1997)などに用いられ、目覚ましい発展を遂げている。遺伝子発現変化のモニタリングと遺伝子ノックアウト技術の組み合わせ等により、シグナル伝達反応経路や遺伝子のネットワークを解明しようという、理論的研究も行われている(Thieffry, 1999; Akutsu et al., 1999)。計算量・比較法など解決しなければならない問題は残されているが、このような研究から、遺伝子が機能する生化学的反応経路、発現時期及び場所といった遺伝子の機能するシステム全体を、異なる生物種間で包括的に比較することが出来るようになるかもしれない。

本論文のテーマである「生物の進化の分子的理解」は、分子進化学に残された最大のテーマだと述べた。しかしまた、分子レベルの様々な調節機構が、どのように表現型に関係するかを知ることは、現代生命科学の重要課題の一つでもある。本論文では、生物進化の分子的基盤を理解することに挑戦し、その一端を明らかにすることができた。今後、分子進化学、古生物学、発生生物学、分子生物学、情報生物学など多岐に渡る分野の成果が統合され、「生物の進化の分子的理解」が進むものと期待する。

謝辞

本研究を行うにあたり、宮田隆教授には懇切丁寧な指導を賜りました。有意義な助言により、私の研究をここまで導いて下さったことに心より感謝致します。

隈啓一博士、岩部直之博士には研究全般に渡り指導を賜り、有意義な議論をして頂きました。お二人に深く感謝致します。

菅裕博士には、多くの助言を頂きました。また、加藤和貴氏、星山大介氏、山口武志氏、工楽樹洋氏、佐々木剛氏をはじめとする研究室の皆様には、たくさんの議論に応じて頂きました。皆様に心から感謝致します。

渡部洋子博士、窪川かおる博士には、生物試料をころよく提供して頂きました。お礼申し上げます。

最後に、ここに至るまで研究生活を支え続けて下さった友人、家族、そして同僚であり夫である小柳光正さんに深くお礼申し上げます。

文献

- Adachi J, Hasegawa M (1996) Computer science monographs, No. 28, MOLPHY version 2.3: Programs for molecular phylogenetics based on maximum likelihood. The Institute of Statistical Mathematics, Tokyo
- Adam RD (1991) The biology of *Giardia* spp. Microbiol Rev 55:706-732
- Akutsu T, Miyano S, Kuhara S (1999) Identification of genetic networks from a small number of gene expression patterns under the Boolean network model. Pac Symp Biocomput:17-28
- Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD (1994) Molecular biology of the cell. 3rd edition. Garland Publishing, New York
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389-3402
- Andreeva AV, Kutuzov MA (1999) Non-classical protein Ser/Thr phosphatases: what are they for? Biochemistry(Moscow) 64:228-234
- Ayala FJ, Rzhetsky A (1998) Origin of the metazoan phyla: molecular clocks confirm paleontological estimates. Proc Natl Acad Sci U S A 95:606-611
- Barnea G, Silvennoinen O, Shaanan B, Honegger AM, Canoll PD, D'Eustachio P, Morse B, Levy JB, Laforgia S, Huebner K; et al. (1993) Identification of a carbonic anhydrase-like domain in the extracellular region of RPTP gamma defines a new subfamily of receptor tyrosine phosphatases. Mol Cell Biol 13:1497-1506
- Bignon JS, Siminovitch KA (1994) Identification of PTP1C mutation as the genetic defect in motheaten and viable motheaten mice: a step toward defining the roles of protein tyrosine phosphatases in the regulation of hemopoietic cell differentiation and function. Clin Immunol Immunopathol 73:168-179
- Bloomquist BT, Darlington DN, Mains RE, Eipper BA (1994) RESP18, a novel endocrine secretory protein transcript, and four other transcripts are regulated in parallel with pro-opiomelanocortin in melanotropes. J Biol Chem 269:9113-9122
- Brady-Kalnay SM, Tonks NK (1995) Protein tyrosine phosphatases as adhesion receptors. Curr Opin Cell Biol 7:650-657
- Bromham L, Rambaut A, Fortey R, Cooper A, Penny D (1998) Testing the Cambrian explosion hypothesis by using a molecular dating technique. Proc Natl Acad Sci U S A 95:12386-12389
- Cavalier-Smith T (1989) Molecular phylogeny. Archaeobacteria and Archezoa. Nature 339:100-01
- Cavalier-Smith T (1993) Kingdom protozoa and its 18 phyla. Microbiol Rev 57:953-994
- Charbonneau H, Tonks NK (1992) 1002 protein phosphatases? Annu Rev Cell Biol 8: 463-493
- Chien CB (1996) PY in the fly receptor-like tyrosine phosphatases in axonal pathfinding. Neuron 16:1065-1068

- Clayton RA, White O, Ketchum KA, Venter JC (1997) The first genome from the third domain of life. *Nature* 387:459-462
- Cohen PT (1997) Novel protein serine/threonine phosphatases: variety is the spice of life. *Trends Biochem Sci* 22:245-251
- Colbert EH, Morales M (1991) *Evolution of the vertebrates*. Wiley-Liss, New York
- Condeelis J (1995) Elongation factor 1 α , translation and the cytoskeleton. *Trends Biochem Sci* 20:169-170
- Conway Morris S (1993) The fossil record and the early evolution of the Metazoa. *Nature* 361:219-225
- Dayhoff MO (1978) Survey of new data and computer methods of analysis. In: Dayhoff MO (ed) *Atlas of protein sequence and structure*. National Biochemical Research Foundation, Washington DC
- Denu JM, Stuckey JA, Saper MA, Dixon JE (1996) Form and function in protein dephosphorylation. *Cell* 87:361-364
- DeRisi JL, Iyer VR, Brown PO (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278:680-686
- Dickerson RE (1971) The structures of cytochrome c and the rates of molecular evolution. *J Mol Evol* 1:26-45
- Doolittle WF, Brown JR (1994) Tempo, mode, the progenote, and the universal root. *Proc Natl Acad Sci U S A* 91:6721-6728
- Doolittle RF, Feng DF, Tsang S, Cho G, Little E (1996) Determining divergence times of the major kingdoms of living organisms with a protein clock. *Science* 271:470-477
- Eizinger A, Jungblut B, Sommer RJ (1999) Evolutionary change in the functional specificity of genes. *Trends Genet* 15:197-202
- Faux MC, Scott JD (1996) More on target with protein phosphorylation: conferring specificity by location. *Trends Biochem Sci* 21:312-315
- Felsenstein J (1978) Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology* 27:401-410
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783-791
- Feng DF, Doolittle RF (1987) Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J Mol Evol* 25:351-360
- Forsyth IA (1994) Comparative aspects of placental lactogens: structure and function. *Exp Clin Endocrinol* 102:244-251
- Frohman MA, Dush MK, Martin GR (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc Natl Acad Sci U S A* 85:8998-9002
- Futuyma DJ (1998) *Evolutionary Biology*, 3rd ed. Sinauer Associates, Inc., Sunderland, Massachusetts.
- Ganformina MD, Sanchez D (1999) Generation of evolutionary novelty by functional shift. *Bioessays* 21:432-439
- Gillin FD, Reiner DS, McCaffery JM (1996) Cell biology of the primitive eukaryote *Giardia lamblia*. *Annu Rev Microbiol* 50:679-705

- Grant PR (1986) Ecology and evolution of Darwin's finches. Princeton University Press, Princeton, NJ
- Han TM, Runnegar B (1992) Megascopic eukaryotic algae from the 2.1-billion-year-old neogranite iron-formation, Michigan. *Science* 257:232-235
- Harrison FW, Cowden RR (1982) Developmental biology of freshwater invertebrates. Alan R. Liss, Inc., New York
- Hasegawa M, Fujiwara M (1993) Relative efficiencies of the maximum likelihood, maximum parsimony, and neighbor-joining methods for estimating protein phylogeny. *Mol Phylogenet Evol* 2:1-5
- Hashimoto T, Nakamura Y, Nakamura F, Shirakura T, Adachi J, Goto N, Okamoto K, Hasegawa M (1994) Protein phylogeny gives a robust estimation for early divergences of eukaryotes: phylogenetic place of a mitochondria-lacking protozoan, *Giardia lamblia*. *Mol Biol Evol* 11:65-71
- Hashimoto T, Sanchez LB, Shirakura T, Muller M, Hasegawa M (1998) Secondary absence of mitochondria in *Giardia lamblia* and *Trichomonas vaginalis* revealed by valyl-tRNA synthetase phylogeny. *Proc Natl Acad Sci U S A* 95:6860-6865
- Holland PW, Garcia-Fernandez J (1996) Hox genes and chordate evolution. *Dev Biol* 173:382-395
- Hoshiyama D, Suga H, Iwabe N, Koyanagi M, Nikoh N, Kuma K, Matsuda F, Honjo T, Miyata T (1998) Sponge *Pax* cDNA related to *Pax-2/5/8* and ancient gene duplications in the *Pax* family. *J Mol Evol* 47:640-648
- Hubbard MJ, Cohen P (1993) On target with a new mechanism for the regulation of protein phosphorylation. *Trends Biochem Sci* 18:172-177
- Hunter T (1995) Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell* 80:225-236
- Iwabe N, Kuma K, Hasegawa M, Osawa S, Miyata T (1989) Evolutionary relationship of archaeobacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proc Natl Acad Sci U S A* 86:9355-9359
- Iwabe N, Kuma K, Miyata T (1996) Evolution of gene families and relationship with organismal evolution: rapid divergence of tissue-specific genes in the early evolution of chordates. *Mol Biol Evol* 13:483-493
- James P (1997) Of genomes and proteomes. *Biochem Biophys Res Commun* 231:1-6
- Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* 8:275-282
- Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) *Mammalian Protein Metabolism III*. Academic Press, New York
- Kasahara M, Hayashi M, Tanaka K, Inoko H, Sugaya K, Ikemura T, Ishibashi T (1996) Chromosomal localization of the proteasome Z subunit gene reveals an ancient chromosomal duplication involving the major histocompatibility complex. *Proc Natl Acad Sci U S A* 93:9096-9101

- Katoh K, Miyata T A heuristic approach of maximum likelihood method for inferring phylogenetic tree and an application to the mammalian SOX-3 origin of the testis-determining gene SRY. FEBS Lett in press
- Kimura M (1983) The neutral theory of molecular evolution. Cambridge University Press, Cambridge
- Kimura M (1991) Recent development of the neutral theory viewed from the Wrightian tradition of theoretical population genetics. Proc Natl Acad Sci U S A 88:5969-5973
- Kishino H, Miyata T, Hasegawa M (1990) Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. J Mol Evol 30:151-161
- Kobayashi M, Takahashi M, Wada H, Satoh N (1993) Molecular phylogeny inferred from sequences of small subunit ribosomal DNA, supports the monophyly of the metazoa. Zoolog Sci 10:827-833
- Koyanagi M, Suga H, Hoshiyama D, Ono K, Iwabe N, Kuma K, Miyata T (1998a) Ancient gene duplication and domain shuffling in the animal cyclic nucleotide phosphodiesterase family. FEBS Lett 436:323-328
- Koyanagi M, Ono K, Suga H, Iwabe N, Miyata T (1998b) Phospholipase C cDNAs from sponge and hydra: antiquity of genes involved in the inositol phospholipid signaling pathway. FEBS Lett 439:66-70
- Krueger NX, Saito H (1992) A human transmembrane protein-tyrosine-phosphatase, PTP zeta, is expressed in brain and has an N-terminal receptor domain homologous to carbonic anhydrases. Proc Natl Acad Sci U S A 89:7417-7421
- Kuma K, Iwabe N, Miyata T (1995a) Functional constraints against variations on molecules from the tissue level: slowly evolving brain-specific genes demonstrated by protein kinase and immunoglobulin supergene families. Mol Biol Evol 12:123-130
- Kuma K, Nikoh N, Iwabe N, Miyata T (1995b) Phylogenetic position of *Dictyostelium* inferred from multiple protein data sets. J Mol Evol 41:238-246
- Lander ES (1999) Array of hope. Nat Genet 21:3-4
- Lee J, Chen Y, Tolstykh T, Stock J (1996) A specific protein carboxyl methyltransferase that demethylates phosphoprotein phosphatase 2A in bovine brain. Proc Natl Acad Sci U S A 93:6043-6047
- Leng J, Cameron AJ, Buckel S, Kennelly PJ (1995) Isolation and cloning of a protein-serine/threonine phosphatase from an archaeon. J Bacteriol 177:6510-6517
- Lundin LG (1993) Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and the house mouse. Genomics 16:1-19
- Mansouri A, Hallonet M, Gruss P (1996) Pax genes and their roles in cell differentiation and development. Curr Opin Cell Biol 8:851-857
- Margulis L, Schwartz KV (1998) Five Kingdoms: an illustrated guide to the phyla of life on earth, 3rd ed. W. H. Freeman and Company, New York
- Matthews RJ, Bowne DB, Flores E, Thomas ML (1992) Characterization of hematopoietic intracellular protein tyrosine phosphatases: description of a phosphatase containing an SH2 domain and another enriched in proline-, glutamic acid-, serine-, and threonine- rich sequences. Mol Cell Biol 12:2396-2405

- Mauro LJ, Dixon JE (1994) 'Zip codes' direct intracellular protein tyrosine phosphatases to the correct cellular 'address'. *Trends Biochem Sci* 19:151-155
- Miyata T, Yasunaga T (1980) Molecular evolution of mRNA: a method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application. *J Mol Evol* 16:23-36
- Miyata T, Kuma K, Iwabe N, Nikoh N (1994) A possible link between molecular evolution and tissue evolution demonstrated by tissue specific genes. *Jpn J Genet* 69:473-480
- Mumby MC, Walter G (1993) Protein serine/threonine phosphatases: structure, regulation, and functions in cell growth. *Physiol Rev* 73:673-699
- Needleman SB, Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* 48:443-453
- Nikoh N, Hayase N, Iwabe N, Kuma K, Miyata T (1994) Phylogenetic relationship of the kingdoms Animalia, Plantae, and Fungi, inferred from 23 different protein species. *Mol Biol Evol* 11:762-768
- Nikoh N, Iwabe N, Kuma K, Ohno M, Sugiyama T, Watanabe Y, Yasui K, Shi-cui Z, Hori K, Shimura Y, Miyata T (1997) An estimate of divergence time of Parazoa and Eumetazoa and that of Cephalochordata and Vertebrata by aldolase and triose phosphate isomerase clocks. *J Mol Evol* 45:97-106
- Ohno S (1970) *Evolution by gene duplication*. Springer Verlag, New York
- Ohta T (1991) Multigene families and the evolution of complexity. *J Mol Evol* 33:34-41
- Ottlie S, Chernoff J, Hannig G, Hoffman CS, Erikson RL (1992) The fission yeast genes *pyp1+* and *pyp2+* encode protein tyrosine phosphatases that negatively regulate mitosis. *Mol Cell Biol* 12:5571-5580
- Palopoli MF, Patel NH (1998) Evolution of the interaction between Hox genes and a downstream target. *Curr Biol* 8:587-590
- Patel NH, Martin-Blanco E, Coleman KG, Poole SJ, Ellis MC, Kornberg TB, Goodman CS (1989) Expression of *engrailed* proteins in arthropods, annelids, and chordates. *Cell* 58:955-968
- Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. *Proc Natl Acad Sci U S A* 85:2444-2448
- Pebusque MJ, Coulier F, Birnbaum D, Pontarotti P (1998) Ancient large-scale genome duplications: phylogenetic and linkage analyses shed light on chordate genome evolution. *Mol Biol Evol* 15:1145-1159
- Perou CM, Jeffrey SS, van de Rijn M, Rees CA, Eisen MB, Ross DT, Pergamenschikov A, Williams CF, Zhu SX, Lee JC, Lashkari D, Shalon D, Brown PO, Botstein D (1999) Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc Natl Acad Sci U S A* 96:9212-9217
- Postlethwait JH, Yan YL, Gates MA, Horne S, Amores A, Brownlie A, Donovan A, Egan ES, Force A, Gong Z, Goutel C, Fritz A, Kelsh R, Knapik E, Liao E, Paw B, Ransom D, Singer A, Thomson M, Abduljabbar TS, Yelick P, Beier D, Joly JS, Larhammar D, Rosa F, et al. (1998) Vertebrate genome evolution and the zebrafish gene map. *Nat Genet* 18:345-349

- Pulido R, Serra-Pages C, Tang M, Streuli M (1995) The LAR/PTP delta/PTP sigma subfamily of transmembrane protein-tyrosine- phosphatases: multiple human LAR, PTP delta, and PTP sigma isoforms are expressed in a tissue-specific manner and associate with the LAR- interacting protein LIP.1. *Proc Natl Acad Sci U S A* 92:11686-11690
- Rasmont R (1961) Une technique de culture des eponges d'eau douce en milieu controle. *Ann Soc R Zool Belg* 91:147-156
- Roberts RM, Liu L, Guo Q, Leaman D, Bixby J (1998) The evolution of the type I interferons. *J Interferon Cytokine Res* 18:805-816
- Roger AJ, Svard SG, Tovar J, Clark CG, Smith MW, Gillin FD, Sogin ML (1998) A mitochondrial-like chaperonin 60 gene in *Giardia lamblia*: evidence that diplomonads once harbored an endosymbiont related to the progenitor of mitochondria. *Proc Natl Acad Sci U S A* 95:229-234
- Rogers S, Wells R, Rechsteiner M (1986) Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* 234:364-368
- Romer AS, Parsons TS (1986) *The vertebrate body*, 6th ed. Saunders College Pub, Philadelphia
- Rutherford SL, Lindquist S (1998) Hsp90 as a capacitor for morphological evolution. *Nature* 396:336-342
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406-425
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, New York
- Sap J, Jiang YP, Friedlander D, Grumet M, Schlessinger J (1994) Receptor tyrosine phosphatase R-PTP-kappa mediates homophilic binding. *Mol Cell Biol* 14:1-9
- Saxton TM, Henkemeyer M, Gasca S, Shen R, Rossi DJ, Shalaby F, Feng GS, Pawson T (1997) Abnormal mesoderm patterning in mouse embryos mutant for the SH2 tyrosine phosphatase Shp-2. *Embo J* 16:2352-2364
- Schaapveld R, Wieringa B, Hendriks W (1997) Receptor-like protein tyrosine phosphatases: alike and yet so different. *Mol Biol Rep* 24:247-262
- Seimiya M, Ishiguro H, Miura K, Watanabe Y, Kurosawa Y (1994) Homeobox-containing genes in the most primitive metazoa, the sponges. *Eur J Biochem* 221:219-225
- Shenolikar S (1994) Protein serine/threonine phosphatases--new avenues for cell regulation. *Annu Rev Cell Biol* 10:55-86
- Sidow A (1992) Diversification of the Wnt gene family on the ancestral lineage of vertebrates. *Proc Natl Acad Sci U S A* 89:5098-5102
- Smith MW, Aley SB, Sogin M, Gillin FD, Evans GA (1998) Sequence survey of the *Giardia lamblia* genome. *Mol Biochem Parasitol* 95:267-280
- Sogin ML, Gunderson JH, Elwood HJ, Alonso RA, Peattie DA (1989) Phylogenetic meaning of the kingdom concept: an unusual ribosomal RNA from *Giardia lamblia*. *Science* 243:75-77
- Stoker AW (1996) Axon guidance: motor-way madness. *Curr Biol* 6:794-797

- Stoker A, Dutta R (1998) Protein tyrosine phosphatases and neural development. *Bioessays* 20:463-472
- Strachan T, Read AP (1994) PAX genes. *Curr Opin Genet Dev* 4:427-438
- Streuli M (1996) Protein tyrosine phosphatases in signaling. *Curr Opin Cell Biol* 8:182-188
- Suga H, Koyanagi M, Hoshiyama D, Ono K, Iwabe N, Kuma K, Miyata T (1999a) Extensive gene duplication in the early evolution of animals before the parazoan-eumetazoan split demonstrated by G proteins and protein tyrosine kinases from sponge and hydra. *J Mol Evol* 48:646-653
- Suga H, Hoshiyama D, Kuraku S, Katoh K, Kubokawa K, Miyata T (1999b) Protein tyrosine kinase cDNAs from amphioxus, hagfish and lamprey: isoform duplications around the divergence of cyclostomes and gnathostomes. *J Mol Evol* 49:601-608
- Swaffield JC, Purugganan MD (1997) The evolution of the conserved ATPase domain (CAD): reconstructing the history of an ancient protein module. *J Mol Evol* 45:549-563
- Swofford DL, Olsen GJ, Waddell PJ, Hillis DM (1996) Phylogenetic inference. In: Hillis DM, Moritz C, Mable BK (eds) *Molecular Systematics*, 2nd edition. Sinauer Associates, Sunderland, Massachusetts
- Taylor SS, Buechler JA, Yonemoto W (1990) cAMP-dependent protein kinase: framework for a diverse family of regulatory enzymes. *Annu Rev Biochem* 59:971-1005
- Thieffry D (1999) From global expression data to gene networks. *Bioessays* 21:895-899
- Toh H, Hayashida H, Miyata T (1983) Sequence homology between retroviral reverse transcriptase and putative polymerases of hepatitis B virus and cauliflower mosaic virus. *Nature* 305:827-829
- Wada H, Satoh N (1994) Details of the evolutionary history from invertebrates to vertebrates, as deduced from the sequences of 18S rDNA. *Proc Natl Acad Sci U S A* 91:1801-1804
- Wainright PO, Hinkle G, Sogin ML, Stickel SK (1993) Monophyletic origins of the metazoa: an evolutionary link with fungi. *Science* 260:340-342
- Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsie L, Topaloglou T, Hubbell E, Robinson E, Mittmann M, Morris MS, Shen N, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson TJ, Lander ES, et al. (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* 280:1077-1082
- Willmer P (1990) *Invertebrate relationships: pattern in animal evolution*. Cambridge University Press, Cambridge
- Wistow G, Anderson A, Piatigorsky J (1990) Evidence for neutral and selective processes in the recruitment of enzyme-crystallins in avian lenses. *Proc Natl Acad Sci U S A* 87: 6277-6280
- Woese CR, Fox GE (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A* 74:5088-5090
- Wray GA, Levinton JS, Shapiro LH (1996) Molecular evidence for deep precambrian divergences among metazoan phyla. *Science* 274:568-573

- Xu Q, Fu HH, Gupta R, Luan S (1998) Molecular characterization of a tyrosine-specific protein phosphatase encoded by a stress-responsive gene in Arabidopsis. *Plant Cell* 10:849-857
- Zhuo S, Clemens JC, Stone RL, Dixon JE (1994) Mutational analysis of a Ser/Thr phosphatase. Identification of residues important in phosphoesterase substrate binding and catalysis. *J Biol Chem* 269:26234-26238
- Zolnierowicz S, Hemmings BA (1996) Protein phosphatases on the piste. *Trends in cell biology* 6:359-362
- Zuckerkindl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ (eds) *Evolving Genes and Proteins*. Academic Press, New York

岩波生物学辞典 第4版 (1996) 岩波書店

小野香奈子 (1997) 蛋白質のリン酸化・脱リン酸化に関与する超遺伝子族の進化 修士論文

武田誠郎 (1998) プロテインホスファターゼ研究の現状と展望 蛋白質核酸酵素

(1998年6月号増刊) vol.43 No.8:927-934

水野一也 (1998) 神経細胞分化におけるチロシンホスファターゼの役割 蛋白質核酸酵素

(1998年6月号増刊) vol.43 No.8:1136-1143

付録1 RT-PCRによるホスファターゼ遺伝子の単離に用いたプライマーの組み合わせを示す。プライマーについては材料と方法の章を参照。-は最初のPCRで目的のDNA断片の増幅が見られたため、nested PCRを行わなかったことを意味する。AUAPは逆転写の際に用いたアダプタープライマーの5'末端側25ヌクレオチドに相当するプライマーである。

1-1 PTP遺伝子及びPP遺伝子の単離に用いたdegenerate primerの組み合わせ

生物	最初のPCRで用いたプライマー		nested PCRで用いたプライマー		単離された遺伝子
	S	A	S	A	
PTP遺伝子族					
カワカイメン	ptp-1	ptp-a	-	-	sPTPR4, sPTPR5, sPTPN1, sPTPN2, sPTPN3, sPTPN6, sPTPN8
	ptp-1	ptp-b	-	-	sPTPR4, sPTPR2B, sPTPR3, sPTPR7, sPTPN6, sPTPN8
ナメクジウオ	ptp-1	ptp-a	-	-	amPTPR4a, amPTPR2A
	ptp-1	ptp-b	-	-	amPTPR3, amPTPR10, amPTPR4b, amPTPR4c, amPTPR5
	ptp-2	ptp-a	-	-	amPTPR4a, amPTPR4b, amPTPR2A
	ptp-2	ptp-b	-	-	amPTPR3, amPTPR4c
ヌタウナギ	ptp-1	ptp-a	-	-	hgPTPR2Aa
	ptp-1	ptp-b	-	-	hgPTPR5a, hgPTPR2Ab, hgPTPR4
	ptp-2	ptp-a	-	-	hgPTPR2Ab, hgPTPR4, hgPTPR3, hgPTPR2B, hgPTPR2Ac, hgPTPN6a, hgPTPN6b
	ptp-2	ptp-b	-	-	hgPTPR2Aa, hgPTPR5a, hgPTPR2Ab, hgPTPR3, hgPTPR2Ac, hgPTPN6a, hgPTPR5b
	ptp-1	ptp-b	ptp-1	ptp-a	hgPTPR2Aa, hgPTPR2Ac, hgPTPN3
	ptp-1	ptp-b	ptp-2	ptp-a	hgPTPN3
	ptp-2	ptp-a	ptp-1	ptp-a	hgPTPN3
	ptp-2	ptp-b	ptp-2	ptp-a	hgPTPR2Aa
エイ	ptp-1	ptp-a	-	-	ryPTPN6a
	ptp-1	ptp-b	ptp-1	ptp-a	ryPTPN6a
	ptp-2	ptp-b	ptp-1	ptp-a	ryPTPR2Aa
	ptp-2	ptp-b	ptp-1	ptp-b	ryPTPR2Aa

1-1 続き

PP遺伝子族

カワカイメン	pp-1	pp-a	pp-2	pp-a	sPP5, sPP2B, sPP2A, sPP1c, sPP1a, sPP1b
	pp-2	pp-a	pp-3	pp-a	sPP2B, sPP2A, sPP1c, sPP1a
ナメクジウオ	pp-2	pp-a	-	-	amPP2A, amPP1
	pp-2	pp-a	pp-3	pp-a	amPP2B, amPP1
	pp-1	pp-a	pp-2	pp-a	amPP5, amPP1
	pp-2	AUAP	pp-2	pp-a	amPP5
ヌタウナギ	pp-1	pp-a	pp-2	pp-a	hgPP2A, hgPP2B
	pp-2	pp-a	pp-3	pp-a	hgPP2A, hgPP2B, hgPP5, hgPP1a, hgPP1b
エイ	pp-1	pp-a	pp-2	pp-a	ryPP5, ryPP4a, ryPP4b, ryPP1a
	pp-2	pp-a	pp-3	pp-a	ryPP2Ba, ryPP2Bb, ryPP4a, ryPP4b, ryPP2Aa, ryPP2Ab, ryPP1b, ryPP1a
	pp-2	AUAP	pp-2	pp-a	ryPP2Bb, ryPP2Ba, ryPP4a, ryPP2Ab, ryPP1b

Sはセンス側のプライマー、Aはアンチセンス側のプライマーである。

1-2 PTP遺伝子の特定のサブタイプの単離に用いたdegenerate primerの組み合わせ

サブタイプ	生物	最初のPCRで用いたプライマー		nested PCRで用いたプライマー		単離された遺伝子
		S	A	S	A	
PTPR2A	エイ	R2A-1 R2A-1	ptp-a ptp-b	ptp-1 ptp-1	ptp-a ptp-b	ryPTPR2Ac, ryPTPR2Ab ryPTPR2Ab
PTPR4	エイ	ptp-1 ptp-1	R4-a R4-a	ptp-1 ptp-1	ptp-a ptp-b	ryPTPR4a, ryPTPR4b ryPTPR4b
PTPR5	エイ	ptp-1 R5-1 R5-1 R5-2	R5-a R5-a ptp-b R5-a	ptp-1 ptp-1 ptp-1 ptp-1	ptp-a ptp-a ptp-a ptp-a	ryPTPR8 ryPTPR5a, ryPTPR5b ryPTPR5b ryPTPR5a
PTPN3	ナメクジウオ エイ	N3-1 N3-1	N3-a N3-a	ptp-1 ptp-1	ptp-a ptp-a	amPTPN3 ryPTPN3
PTPN6	ナメクジウオ エイ	N6-1 N6-1 N6-1 N6-1	ptp-a N6-a N6-b ptp-a	ptp-1 - - ptp-1	ptp-a - - ptp-a	amPTPN6 ryPTPN6b, ryPTPN6c ryPTPN6c ryPTPR2B

カワカイメンやヌタウナギでは、1-1に示したプライマーの組み合わせを用いることで、本論文で示した全ての遺伝子が単離された。一方ナメクジウオや特にエイでは、1-1のプライマーを用いたPCRでは、少数の遺伝子しか単離できなかった。そこで、特定のサブタイプに特徴的なプライマーを用いることで、単離されなかったサブタイプに属する遺伝子の単離を行った。4種の生物ともできるだけ多種類の遺伝子を単離するように留意した。Sはセンス側のプライマー、Aはアンチセンス側のプライマーである。

付録2 本論文で単離した遺伝子の塩基配列及び推定アミノ酸配列を示す。生物種名、遺伝子名、GenBankのアクセッションナンバーを記し、塩基配列の下に推定アミノ酸配列を一字表記にて示した。・は終止コドンを表す。カワカイメンのsPTPN2遺伝子については、PEST配列検索ソフトウェアPESTfindによりPEST配列である可能性が示された領域を四角で囲った。

2-1 カワカイメンのPTP遺伝子の塩基配列及び推定アミノ酸配列

sponge sPTPR4 (AB019125)

```
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGCAGGGCAGTGCAACAATTGTAATGTTGACAACTTGGAGGAGAAAGGAAGGATTAAATGCAATCAGTATTGGCCAGAAGGAAGAACT
E Q G S A T I V M L T N L E E K G R I K C N Q Y W P E G R T
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCGGTGTATGGATCTATTAGAATTCACACTCAGGAGGAGTTAACTCTCGCCGACTACACATTGCGAATATTCTCACTTCAGCGGGAAGGT
A V Y G S I R I H T Q E E L T L A D Y T L R I F S L Q R E G
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCCAATGAAGAAAGAATAGTGAAACAATTCCATTTACTGTTTGGCCAGACCATGATGTACCTGACTATCCAAGTGCATTGCTTCATTTT
S N E E R I V K Q F H F T V W P D H D V P D Y P T A L L H F
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTTTCGTAAAGTCTCTGCATCCAATACAGAAAAGTCTGGTTCCTATGGTTGTTTCATTGCAGTGCAGGTGTGGGGCGCACTGGGACCTTCATA
V R K V S A S N T E K S G P M V V H C S A G V G R T G T F I
-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACGCTGCATTCCCAAATGAATCGCTTACGTGAAGAAAGAAATATTGATGTGTTTGCTTTTCGTTAAAGGAATGAGATGCCAGAGGTGTCTC
T L H S Q M N R L R E E R N I D V F A F V K G M R C Q R C L
-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATGGTTCAGACTGAGCGTCAATATATCTTTATTCATGATGCCTTACTAGAAGCTATTGAGTGTGGGGACACAGATGTTTCTGCAAGGGAT
M V Q T E R Q Y I F I H D A L L E A I E C G D T D V S A R D
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTAAATGAGCAATACAGGCTGCTGTGTGCACTGAATGACACTACGCAAAAGACTATGCTTCAGATGGAGTTTGAGAAAATTAACACCACA
L N E Q Y R L L C A L N D T T Q K T M L Q M E F E K I N T T
-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATACATCGAAAAGTTGCCAAGGATGTTGGTGTTCCTTCATTACAATAGAGTAAAGAATAGATACACAAATATTGTTCTTTTGACAGTCAT
I H R K V A K D V G V L H Y N R V K N R Y T N I V P F D S H
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGGTGTAGAATGAGCATCATTCCTGGTGTGAGGGCTCTGATTACATCAATGCCAGCTTGATAAATGGTTATCACAAACCTCGTGCCTTT
R C R M S I I P G V E G S D Y I N A S L I N G Y H K P R A F
-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATTGCTTCGCAAGGTCCTCTACCAGATACTGCAGATGACTTTTGGAGAATGGTTTGGGAGCAAAAAGTTCATCAATTGTAATGCTTACT
I A S Q G P L P D T A D D F W R M V W E Q K S S S I V M L T
-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAGGAAAGAGAAGGGGGCAAAGTAAATGCCATCGGTATTGGCCAGAAAGTGGAGCAGAGATGCATGGTCCATTTCAAGTTATATTATAT
K E R E G G K V K C H R Y W P E S G A E M H G P F Q V I L Y
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGTATGCATGAGTATCCGGAGTACACATTAAGAGAGTTCAAGTTGATAGACAGCAGGGATCGTGAAGGCAGTATACAATTTGTGAAGCAA
G M H E Y P E Y T L R E F K L I D S R D R E G S I Q F V K Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TACCAGTTTACAGTTTGGCCTGACGAAGGCATTCCAGGGAATGGTGTGGGAATTGTGGATCTCATTGGTCAAGTACAAAAATGGCAGCGC
Y Q F T V W P D E G I P G N G V G I V D L I G Q V Q K W Q R
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGTGTGGAGACAAGCCAATAGTTGTGCATTGCAGTGCTGGATGTGGTTCGCACTGGTACATATGTTGCTATTAGCATTCTGCTGGAAAGG
Q C G D K P I V V H C S A G C G R T G T Y V A I S I L L E R
```

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTGAGGACTGAAGGTGCAGTGGACGTTTTTTCACACTGTACGGACGCTTCGATTGCAACGGGAGTGGATGGTCCAGACAGAGCTGCAATAT
L R T E G A V D V F H T V R T L R L Q R E W M V Q T E L Q Y
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAATTTATTTATAAACTATATTGGAGTACCTGGACTCATTTGAACTGTATTCAAATTTTAAGTAGATGGTTTTTTAT
E F I Y K T I L E Y L D S F E L Y S N F K .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

sponge sPTPR2B (AB019126)

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAATATGACTGCCCAACCATCGTCATGGTGACCAACCTTCAGGAGGAAGACAAGGTCAAGTGTCTCAGTACTGGCCATCATATGGCTCC
E Y D C P T I V M V T N L Q E E D K V K C H Q Y W P S Y G S
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACCACCTATGGTCACCTCAAGGTAAGTCTGAAGGAGGTAGAGAAGTGGCCGAGTACAGCATCCGTACCTTCAGTGTACACCAATATCT
T T Y G H L K V T L K E V E N L A E Y S I R T F S V T P I S
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTCAAAGCCGTTCCCCGATGCCTGTGCGAGAGGTGCGTCAGTTCCACTTCCTGGTGTGGCCGACCATGGAGTGCCCCAGTATGCCACT
G Q S R S P M P V R E V R Q F H F L V W P D H G V P Q Y A T
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCCCTGCTCAGCTTCAGGAAGAGGATCTCCAAGTACCACCCCTCAACAGAGGACCTATGGCCGTCCATTGCAGTGCCGGAGTGGGTCTGA
A L L S F R K R I S K Y H P L N R G P M A V H C S A G V G R
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACAGGGACCTTCATTTGCATCGACTATGTCTGAAGCAGATCGAGGGAGAAGCCATTGTGGACATCTTCAACTTTGTGCGTCATATGCGA
T G T F I C I D Y V L K Q I E G E A I V D I F N F V R H M R
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TACAGAAGGAAGTACATGGTCCAGACTTCTGCTCAGTATATCTTCATTCATGATGCCATCCTGGAGTCAGTGTGTGTGGGGACACCTCT
Y R R N Y M V Q T S A Q Y I F I H D A I L E S V L C G D T S
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATACCAGCCTCTGAACTGAAGGATCGCATGGTTGCCATGAAGCAGTTGGATCCCGAATCACAGATCATGAAGATTGAGACTCAGTTTGAG
I P A S E L K D R M V A M K Q L D P E S Q I M K I E T Q F E
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACTTTGAACACTGCCTCTCCTATGGACAAAGACTTTGACTGCTCCATAGCTGAGCTTCCTGACAACCAAAAGAAGAACAGATTCCATGTT
L N T A S P M D K D F D C S I A E L P D N Q K K N R F H V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CACCTTCCATCCAACAAGTTCAGGGTGCCTCTCTCAGGTGATAAGAACCAGTCAGACTACATCTGTGCCAGCTTCTGGATGGTTATCGT
H L P S N N F R V A L S G D K N Q S D Y I C A S F L D G Y R
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGAGAAGCGCCTTCATCGTCACTCAAGCACCAATGGCAGAGACGGTGGGAGACTTCTGGAGGATGGTCTGGGAGCACAAAGTCTGCGGCC
Q R S A F I V T Q A P M A E T V G D F W R M V W E H K S A A
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATCGTGATGTTGGCTCAGCTGGTGAAGATGGGAAGGAGGTGTGTGTACGTTACTGGCCAGAGGAGATGGGTCTCCAGAGGTGTTTGGT
I V M L A Q L V E D G K E V C V R Y W P E E M G R P E V F G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAGTACCAGGTGGAGAAGAGCACAGAGGACCAGAAATGAAAGCACTTACATCATCAGGAAATTCAAGATATGCTCCATTGACAATCCCCAA
K Y Q V E K S T E D R N E S T Y I I R K F K I C S I D N P Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GACAACAGGATGGTCACTCAGTTCCACTACTTGATGTGGCCGAACCGTGCAGTCCCGAGAACACCAATCAATTGTTGAGTTGATTGAT
D N R M V T Q F H Y L M W P N R A A P E N T K S I V E L I D
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```



```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCAAGGCGTATTGACGAGGGCAACTACACTAACAACGCCATCCTTGTGCATGACACTACCAGCCTGGGTCCCTCTGGCATCTTCTGTGCC
A R R I D E G N Y T N N A I L V H D T T S L G P S G I F C A
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTGTGGCACTGCATGGAAAGGCTGAGGTTAGAGAGAGTTGTGGATGTCTTCCAGGCGGTGCAGAAGCTTCAGCTGCAGAAGCCAGGCATG
L W H C M E R L R L E R V V D V F Q A V Q K L Q L Q K P G M
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTGGAGTCGGTCAAGCAATATGCATTCATCTACGACTGCGTTCATGAGTTTATTAATATCACCACCTCCACCACAGTCTGACGTAATATG
V E S V K Q Y A F I Y D C V H E F I N I T T S T T V .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCTCACAACGCTCATTGGTTGTGATCATCAACAAAGTGTGCGTATAATACACTGATCTGCACACA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

sponge sPTPR3 (AB019128)

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGGGACAGCTGCAGACCATTGTGATGCTGACAAAATGTTCTGAAGGTGGCAAGATCAAATGTGAAATGTACTGGTCAGATAGCATGGGA
E G Q L Q T I V M L T K C S E G G K I K C E M Y W S D S M G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGAGTCTATGAGACTGACACGCTCACAGTAACAACCACTTCCAATGAGTCTATGCTGACTTCCAATCCGAAAGTTCTCTATCAAATGT
G V Y E T D T L T V T T T S N E S Y A D F Q I R K F S I K C
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAAGTGGCTCCAGATGCGGACTGCCTCCAAGTGACCCACTTCCACTACACTGCCTGGCCAGACCATGGAGTACCCAGACTGCTACCTCC
K V A P D A D C L Q V T H F H Y T A W P D H G V P Q T A T S
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATCCTCAGCTTTGTGAGGAGGGTACAGAATGCCCATGACAAGGGCAAGGGGTTCCCCTCCTGGTTCAGTGCAGTGTGGGGTGGGGCGT
I L S F V R R V Q N A H D K G K G V P L L V H C S A G V G R
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACGGGAACCTTTCAATTGCACTGGACACACTTCTAGACAGGGTGAGGTGAGAGACATCCATCAGTGTATTTGAGATTGTCAAGGATATGAGA
T G T F I A L D T L L D R V R S E T S I S V F E I V K D M R
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AATAGACGAAGATTTCATGATCCAAACACTGGCACAGTATGTGCTCATCTATGATGCCTTTGATAAGTATATCACGTGTGGGGACACTGGT
N R R R F M I Q T L A Q Y V L I Y D A F D K Y I T C G D T G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATCAGTGAATAATCCTCGATGCTGAGCAGAGATAATAATAACAAATAATGATTATAATTATTCATTTGAGAACGTTGATGTA
S E .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

sponge sPTPN1 (AB019129)

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GACAACACCGTGCTGGTTATCGTCATGACAACAAAGTTAGTTGAGGCAGCCAAAGTGAAGTGTGCAGAATACTGGTCCGACGAGAGGCAG
D N T V L V I V M T T K L V E A A K V K C A E Y W S D E R Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAACAACATGGAGGGGTGATCATTGAGTTACAAAGGTCGAGAAGTTTGAAGGATATGAACAGAGAACAATGCGTGTTTCAGTTTCGAGGGA
E Q H G G V I I Q V T K V E K F E G Y E Q R T M R V Q F E G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAAGAACGGACCCTCGCACACTTCCAATTCTTGGCGTGGCCCCGATTACGGCATTCCCAGCAGCGGTTCCAGCGTCTGAATCTGCTGCGC
E E R T L A H F Q F L A W P D Y G I P S S G S S V L N L L R
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCAGTGCAGGAAACGCAAGCCAAGCAAGCGAAACGGTTCGCCGAATGCCTTTATATGGGCCTCCAGTGTTGGTTCACTGCAGTGCCGGC
A V R E T Q A K Q A K T V A G M P L Y G P P V L V H C S A G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

+-----+-----+-----+-----+-----+-----+-----+-----+
 GTCGGCCGGAGCGGAGCCTTCTGTGCCATCGACTATTGCATCGATGAAGTTTCGCGACCTGGGTCAAGTGAACGTGCAGGGAGCGGTGCGG
 V G R S G A F C A I D Y C I D E L R D L G Q V N V Q G A V R

AAACTAAGGCGTCAGCGGGCCTACGCCATCCAGACGGAGGAACAGTACGTTTTGTGCTACAGAAGTGTCTCGAGTACGCCAAGAGCGTG
K L R R Q R A Y A I Q T E E Q Y V L C Y R T V L E Y A K S V

TGACCGATGACGCGCAAAGAGAGAATTGT

-----+-----+-----+-----+-----+-----+-----+-----+
GAACAAAGGGTGAAATGTGTTGTGATGACTACAAACATCATTGAGAAGGATTCAAAGAAATGTGAAAAGTACTGGCCTGAGCCAGGCCAC
E Q R V K C V V M T T N I I E K D S K K C E K Y W P E P G H

-----+-----+-----+-----+-----+-----+-----+-----+
ACCCTGAGGTTTGGTGCAATCACTGTGAGTGGCATCCAGCAAAAACAAGAGGAAGGCTTCATATTGTCAAACACTCACACTGGCCCATGAC
T L R F G A I T V S G I Q Q K Q E E G F I L S K L T L A H D

-----+-----+-----+-----+-----+-----+-----+-----+
AGCGAACCTCAACAACATCAAGTCTACCACTTCCACTACACACTTTGGCCCGACTTTGGAAGCCCACCTTCGACAGTAGAATTCAATGCC
S E P Q Q H Q V Y H F H Y T L W P D F G S P P S T V E F N A

CTTGTGAAGCAGTTCTCCCTCTTCTCCTCCAACCACCCCATACTGGCCCACTGCAGTGCAGGCCTGGGTGCGACTGGTGTGCTGGTGGCA
L V K Q F S L F S S N H P I L A H C S A G L G R T G V L V A

GTGCACACAGCGCTGGAGTACCACAAGGCCGCCACAAAGTCGACCTGCCCGCCATTGTCCGCCGAATGCGTCGCCAGCGTGGAGGGATG
V H T A L E Y H K A G H K V D L P A I V R R M R R Q R G G M

ATACAGACCCCAGAGCAGTACCAGTTCTGCTACCAGGCCATCGCCGACACACTGGTGTCCAAGCGACTCCAGCGGGTGGGCTCGGAGCCC
I Q T P E Q Y Q F C Y Q A I A D T L V S K R L Q R V G S E P

TCGATTAAGAGTCCCAGGCCAGGAGTGACAGACCGTTCTCGGAGCCCACGCAGGGTCAAGATACACCGAGGACGAAGCGGCATAAAATG
S I K S P Q A R S D R P F S E P T Q G Q D T E R T K R H K M

TATCTGCAGCACAAATCTTCCACCCCCACCTGAAGCCCCGCCCCACTGCTTGTGCAACAGACAGGAAGTTTCTCTTCTCTGAAACCCCTT
Y L Q H N L P P P P E A P P P L L V Q Q T G S F S S S E T L

CTTCACCCCTCCCCTGCTCCCATGTCCCCCCCACCTGAAAATTCTTCAGCCTCTGCCTTGATCTCATCCCCTGGAGCCTCGCCCACTTCT
L H P S P A P M S P P P E N S S A S A L I S S P G A S P T S

CCAGCCAACATGGTAGCCATGACCAACTTGTGCGCTGGCAACACAGCCACACCTTTTGTTTTGTCTTGTGGCCGGCTCACCATCAGCC
P A N M V A M T N L S P G N T A T P F V L S L L A G S P S A

ACACATCTGTTGTACCCTCCCAGGTGCGACCCTCCCCAGTGTCCGACAAAGCCAGTGAAGGCACTGTCCCCATTGAGAGCCCCCGCCT
T T S V V P S Q V D P P P V S D K A S E G T V P I Q S P P P

CCAGTTCCTCACCTCCTCCTCCATTTCACCCATGTCACCACCGGAAACAAGCCCTGTGCCGAGTAAATTACCCTGGTGCTGTTGACC
P S S S P P P P I S P M S P P E T S P V P S K L P L V L L T

CTCTCTCCCCTTACAGACGATGCTGATATGGACGAGGATGATCGTATTCTGTACGAAAGCAGACACCCTCGAGTCAAGCCATACAAAAGA
P P P L T D D A D M D E D D R I L Y E S R H P R V K P Y K R

CAGGAGA AACTCTCTGATCATTGTTCATGACAACCAATGAGGTGGAGAGAGGAAGGAACAAATGCACCCGCTATTGGCCTGATCTTGTATGTC
 Q E N S L I I V M T T N E V E R G R N K C T R Y W P D L D V 90
 TCCAAGCTTTACGGGAAGGTGACGGTTTGTTGTGTGAAGGAGACCACCAAGCCGCACTACGTGTTGAGGGAGTTTGTGGTGAACAGGAAC
 S K L Y G K V T V C C V K E T T K P H Y V L R E F V V N R N 180
 TCTGAGGATGATCCAGCAGAGAAGGAAGAACGGAGGACAATATTCCAGTATCACTTCAAGGCATGGCCTGACCATGGTGTGCCCATGAC
 S E D D P A E K E E R R T I F Q Y H F K A W P D H G V P H D 270
 CCCGGGGCAGTGCTGGGCATCCTGCAAGATGTCAATCTGCAGCAGAAGGAAGTGAAGTGAAGGGGGTGAAGCCTGGGCCCATTGTGGTA
 P G A V L G I L Q D V N L Q Q K E L T D E G V K P G P I V V 360
 CACTGCAGTGCTGGGATTGGGGCAACTGGTACCTTCATTGTTATTGATGTCATCCAAGATCTTATCAACCACCAGGGATGGGATTGTGAA
 H C S A G I G R T G T F I V I D V I Q D L I N H Q G W D C E 450
 ATTGACATACAGAAAACCTATTTCAGCTTGTAAGGAGTCAAAGATCTGGTATGGTGCAGACTGAGCAACAATACAAGTTTGTGTATGTAGCA
 I D I Q K T I Q L V R S Q R S G M V Q T E Q Q Y K F V Y V A 540
 ATACAGCATT TTTGTGGAATCCTATCAAGCTATGCTACAAGGGGGAGGTACTGGTGA CTATGGTAACATGACATTCCCTCTTAAAGAACCA
 I Q H F V E S Y Q A M L Q G G G T G D Y G N M T F P L K E P 630
 GTTCCTGATGACCATACGTATGGAAACGTTTCAAAGATGAAACACAGGAGGATGATTAATTGACCAGGCCATGATTGATTGACTCCTG
 V P D D H T Y G N V S K D E T Q E D D . 720
 ATGGCTGTAATATTACTTCATTGTGTGTATAGGGACACGTAGCTGCACCAC 810

GAACAGAAATGTGTCATCATTCTCATGCTAACCAGCGAGCAGGAGGAGGGTCAGCCTAAATGTCATCGTTATTTCCACCTGATGATCAG
 E Q N V S I I L M L T S E Q E E G Q P K C H R Y F P P D D Q 90
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CAAGGCGGACATGTGCAATTCGAGCAGTATCGCATCACTTTGAAATTTGTAGCACAAAACCTCAGTCACTACAAGATGTTTCAGCCTGAGA
 Q G G H V Q F E Q Y R I T L K F V A Q N S V T T R C F S L R 180
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CATATCCCCAAGCAATCAGGTAAGGGAGATAACTCACCTGCAGTATTCTGAGTGGCCTGATCACGGAATACCAGAAGACCCACAACCGTTC
 H I P S N Q V R E I T H L Q Y S E W P D H G I P E D P Q P F 270
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CTCGACTTTGTTGGTACTGTTCACTCATTGAGGCAACAGTACCACTCAGGGGTTCTGTACTAGTCCATTGCAGTGCAGGGGTTGGGAGG
 L D F V G T V H S L R Q Q Y H S G V P V L V H C S A G V G R 360
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TCTGGTGTGTTTGTCTGATGGATTTCTTATGGCAAAGGTGGACTGTGGAGATGAAATTGACATCGCAAATCACTGCAGCTGCTGCGA
 S G V F V L M D F L M A K V D C G D E I D I A K S L Q L L R 450
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CCAAAGAATGAACCTAGTTTCAGATGGTTGGACAGTACAAATTTTGTACAGTGC AATTGTACACTATATAAAAAAGCCCGCCTTATA
 D Q R M N L V Q M V G Q Y K F L Y S A I V H Y I K R A R L I 540
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TGACCAAAAATTGTCTTGTATTCAACTGTATT 630

[illegible]

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCATCTGACAATGCTCTACCTCCACCTCCACCCTCTAGCCTTCCCCCTGGTCCCCCTGTCACTCACTCACCAGAAAGTAGTGACGAAGAG
P S D N A L P P P P P S S L P P G P P V T H S P E S S D E E
810
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAAGAAGAGGAGGAGGTGAGCACAGAAGAAAGGGCAGAAGAGTCCCAACCAAAGAAGTTGAAATTGGATCCACCGCAAACCCCGAACGTA
E E E E E V S T E E R A E E S Q P K K L K L D P P Q T P N V
900
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAAAATGGACACTTGGACAGTGATGATGAGGTCTCTCCACCCCCTCCTCCTATTCCACCTCGCTCACAAAGCCTTTACCTGAGCACACT
E N G H L D S D D E V S P P P P P I P P R S Q S L S P E H T
990
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGTAAACAAAGCCCTTAGCACAATCCAACGATAGCTCGGGTCATGTTTTTTTGTCCCCTCAAGGTGAAAAGCCTTCTGCCTCCTCCTCA
S K T K P L A Q S N D S S G H V F L S P Q G E K P S A S S S
1080
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCACTTGAGCAGGAAGAGGCCCCACCGATATTACCTCCAAGGGTATCAAACCAACCTATTCCACGTCCCCAAGGACATCTGAAAGGGAA
S L E Q E E A P P I L P P K G I K P T Y S T S P R T S E R E
1170
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACCCATTCAGAGGAGAGGCACAACCCAGACACACAGTCACCCACATTTGCAACAAGCGACACCTCGGAGACCATCCAAAGGAAGAAGCAA
T H S E E R H N P D T Q S P T F A T S D T S E T I Q R K K Q
1260
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGAGAGAAGAGCGTCAGAACCGAATAGCCAGCAGAGTAACAGAGATGAAGGAAAGACTGAAGGAGACTGATGCTTCATCATCCCCAGAG
E R E E R Q N R I A S R V T E M K E R L K E T D A S S S P E
1350
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGGTCTGCCCCATTAATGGTTGTGGCTTTTGGTGTGGTTGTTGGCTACTTTGCTTATGCTGTCATATACTCTTAAAGACATGTTGCTAAG
W S A P L M V V A F G V V V G Y F A Y A V I Y S .
1440
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATATAAAAAAAAAA
1530

```

2-2 ナメクジウオ・スタウナギ・エイのPTP遺伝子の塩基配列及び推定アミノ酸配列

amphioxus amPTPR3 (AB033560)

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAACAGAACAGCAGGGTCATCGTCATGCTGACACAGTGCTGGGAGAGGGGAAAGCCTAAATGTGAGCGTTACTGGCCTGAGGACGAGGAG
E Q N S R V I V M L T Q C W E R G K P K C E R Y W P E D E E
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCAGTGTTCTACGGAGACATCGTGGTGAAGATGCTGTCAGAAAACAAGGAAGATGACTGGACATGCAGGGAATTTGAGCTGGCCAATAAC
P V F Y G D I V V K M L S E N K E D D W T C R E F E L A N N
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGCCACAGTCGGTGCGTCCGGCAGTACCAGTTCACCAGCTGGCCTGACCACGGGGTTCCCGAGGACACCACGTCCAGTCTGCAGTTTGTCT
S H S R C V R Q Y Q F T S W P D H G V P E D T T S S L Q F V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGGATGGTCAGGAGGAACATCGGAGAGGGCGCAGGGCCGACTGTGGTCCACTGCAGTGCAGGTGTAGGCAGGAGCGGTACTTTTCACTA
R M V R R N I G E G A G P T V V H C S A G V G R S G T F I T
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTGGACCGGCTGCTCCAGCATATGGAGGACCACGAGCAGGTGGACATCTTCGGCATCGTCCACCAGATGAGGATGCACCGGGTCTTCATG
L D R L L Q H M E D H E Q V D I F G I V H Q M R M H R V F M
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTCCAGACGCAGGAGCAGTATATCTTCATCCACCAGTGCATCATGGACCTTCTCCAGGAGAGCACGCCGGTGAAGCGCCGCTCCACCATC
V Q T Q E Q Y I F I H Q C I M D L L Q E S T P V K R R S T I
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAGAAGGTCCAGTTTGCCGAGGCTGAGGATGACGAAGAGGTGGATCTGGATGACATCAACTTTGCGGAGTCCACGGAATACCTGGTCTCT
K K V Q F A E A E D D E E V D L D D I N F A E S T E Y L V S
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCTGATGCCAACTCTGCCCTCTGATCTTCTTATCTGTACATGTATCTGTATAGCCGCCTTTTCGTGTAACACCCTTAGTCTAGCTTCTGT
P D A N S A L .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATCTGTATAGCCGGTATGTGTGGCGTCGGTGTGGCGCAACGGTTAGAGCGTTAGACTCGGAACCCATGGGTCCCGGGTTTGATACCCACC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATGCCCCGACGTTGTGCCCTTGGGAAAGGCACTTTACACGACTTTCCCTGACGGTGGAGTGACGCCACCAAGCCTCTTAGGCTAATCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TAAAGTGTGCCAAAAACACGTAGCCGGTATAACTGCCCTTCAAGGCTACTGTACAGTGTTCAAAATACTTTTCTGAGTGTGCTACCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGGGCAAAAACCTAGGTAGCGCATAAGAATTTTAGGTTGACGCCGTCGGTCCCAGATCAAAGGGCTGTGCCAGGCGCCACCGTCAATCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGATCTG

```

amphioxus amPTP10 (AB033561)

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGGAGAGGAGCGCCTGCGTCGTCATGGTGACCAACCTAGTGGAAAATGGAAGGGCTAAGTGTGAGCGTTACTGGCCAGAGGAGAACAGT
Q E R S A C V V M V T N L V E N G R A K C E R Y W P E E N S
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGTACGAGAACAACGTGCAGACGTACGGGGACATCACGGTCACGGCGGAGAAGGTCAGCACTATGGCGGACTACGCGCTCAGGCTGCTG
E Y E N N V Q T Y G D I T V T A E K V S T M A D Y A L R L L

```

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+
CACATCCAAAAGGCTGGAGTGGACGAGGTACGGGAGGTGCTACACTTCCAGTACACTAGCTGGCCGGATTATGGGGTACCCAAGCACCCC
H I Q K A G V D E V R E V L H F Q Y T S W P D Y G V P K H P
-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACGTCAACCATCAACTTCGTCAAGAGGGTCAAGGCGTCCATTCTCGCGGGGCTGGACCGACCATTGTGCACTGCAGTGCGGGTGTGGGG
T S T I N F V K R V K A S I P R G A G P T I V H C S A G V G
-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGATCCGGTACCTTCATCACCATCGCCGCCATGTTGGACATGATCGAGGAAGAGGGCGCCATTAATATACACGACTTCGTAGATGCCATG
R S G T F I T I A A M L D M I E E G A I N I H D F V D A M
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGCGAGAACAGGATGACTATGGTGCAGACTTCGGACCACTACGCGTTTATCTACACGGCGCTGCTGGAGGCCACACAGTGCGGCAACACC
R E N R M T M V Q T S D Q Y A F I Y T A L L E A T Q C G N T
-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACCATTCCTGCCGACGAATTCCGCGACAGATTCCAGCACTCTCAGCCGTAAAAACCCCAAACTGGCTTGAATGCGTACCAAGAGGAGTTC
T I P A D E F R D R F S T L S R K N P K T G L N A Y Q E E F
-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAGAGACTGGCTCGGGTCTGTCCGCCCCTGACCCAGACAAGACCAGGTCCGGCCCTGTACCCGGCGAACAAAAACAAGAACCGCTTCTCT
K R L A R V C P P L T P D K T R S A L S P A N K N K N R F S
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCAGTCCTTGCCAGTGACAGTAACCGCCCCATGCTGACCCCCAGGCCCGGCGACCCGCTGACGCCGACTACATCAACGCACTTCCTG
S V L A S D S N R P M L T P R P G D P P D A D Y I N A D F L
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GATGGCTACAAGCGGAAGGACGCGTTCCTCGTCACCCAATCACCGCTGCCCGGACCGTCACTGACTTCTGGCGCATGCTCAGTGATCAC
D G Y K R K D A F L V T Q S P L P G T V T D F W R M L S D H
-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGATCAGTCTCCGTAATCATGCTGAACCACTAAACCCGAAGGACAAGAGCGTTGTCAAGTACTGGCCAGATGAGGGGTCAAAAACCTTTC
R S V S V I M L N Q L N P K D K S V V K Y W P D E G S K T F
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGCAATCTTGAGGTGGAGATGGTGTTCGGAGAGTCGAGAAAAGAGCGCCACCGTGCGGCAGTTCAAGTTGTCTCTCAGACACACGGTTCGAT
G N L E V E M V S E S R E K S A T V R Q F K L S S D T R F D
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGGCAGTCGGACATGTTTACCAGTTCGAGTTCACGGCGTGGCCGAGCGGTCTGTTCAACTCCGACCTCCGTGGACGATTTGTTAGAACTG
R Q S D I V H Q F E F T A W P S G R S T P T S V D D L L E L
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTGGGATTGGTGGAGAAGTGGCAGCAGCAATCTGGCAACGGACCAATCACAGTCCACTGCATTAATGGAGTGGGCCGGAGCGGTGTTTTT
L G L V E K W Q Q Q S G N G P I T V H C I N G V G R S G V F
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGCACTGCGCATGCGATGTTAGACAGGCTAAAGACAGAGCAGGTTGTTGACGTGTTCCAAGCAGTGAAGACTCTGAGGAACCACCGGCCT
C T A H A M L D R L K T E Q V V D V F Q A V K T L R N H R P
-----+-----+-----+-----+-----+-----+-----+-----+-----+
AACATGGTAGAGACCGAGGAGCAGTACCAATTCTGCTACCAGGTGGCGCTTGCGTACCTGGCCTCCTTCGACACTTATGCAAACTTCAAG
N M V E T E E Q Y Q F C Y Q V A L A Y L A S F D T Y A N F K
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TAAGGTGCCAAAATATAGGCAGAGCTTCAGATGTCTTTCCATGAGAGTTTGTACGGCAATGTTGCTAAAATGGATACAGTTTAAGGAATC
.
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TAGTTGCCAGTGTTCAAAATATT

```



```

-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGCTACCAGGCGGTCGTGGAGTACCTTGACAGTTTTGACCACTACTCAAACCTTGCCTAGATATATTCTAGATTGTGGAGAAGACTTCAC
C Y Q A V V E Y L D S F D H Y S N F A .
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TATATACAATGTACCTAGCAAAATTGACACAATGCTATATGTTGCCTTGAAAGTTCAGTTGTGTTGCGTAGAATCCATTAAAAGAAGACT
-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATTTAAATAGGATTGTTAAGACAATATCTTTGTCAAGTAATAAAGCCACCGCTTCTATAACATCAGTGTATATAAATTACACACATGTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACTCAGCAAACTCTCCAAGGAAAGGTTTAGGCTTCGGCTCATTCTTGACGTGTTTTAGTCGTTTTTACT
-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

amphioxus amPTPR4c (AB033564)

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAACAGAACACGGCAACCATTGTCATGGTAACCAATGTCAAGGAGAAGAACAAGGTGAAGTGCACCCAGTACTGGCCAGACACAGGGTCC
E Q N T A T I V M V T N V K E K N K V K C T Q Y W P D T G S
-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAGCAGTACGGTGACATCACGGTGAGGAGTGAGGAGACCTCCACCCTAGTGGATTACATCATCAGGACTCTCACTCTTCAGAGGGGTGAT
K Q Y G D I T V R S E E T S T L V D Y I I R T L T L Q R G D
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGACCCGTACTCTTCTCCACTTCCACTTCACCACCTGGCCGGACTTTGGTGTGCCCAAGAGCCCGCTGGGGATGATGAAGTTTGTGAGG
E T R T L L H F H F T T W P D F G V P K S P L G M M K F V R
-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGAGTGAAGGCTGCCAACCTGCAGACTCAGGACCTATAGTGGTGCAGTGTAGCGCGGGAGTTGGACGGACAGGGACATTTATCGTGATT
R V K A A N P A D S G P I V V H C S A G V G R T G T F I V I
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GACGCCATGTTTGAATGATCGCGGCGGAACAGAGGGTGGACGTGTTTGGATTGTGCGGACAGATCCGACAGAGCAGATGCATGATGGTC
D A M F E M I A A E Q R V D V F G F V G Q I R Q S R C M M V
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGACAGAGGGCCAGTATGTGTTTATTTACCAAGCACTGCTGGAACATTTCTGTACGGTGATACAGAGATTGAGGTGACAAACCTGCGC
Q T E G Q Y V F I Y Q A L L E H F L Y G D T E I E V T N L R
-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGACACCTGCAGCAGCTGGCGGCCAAGCTGCCGGACTCTCAGGACACCGGCATGGAGGCTGAGTTCAAGAACTGACCCAGATCCCCATT
H L Q Q L A A K L P D S Q D T G M E A E F K K L T Q I P I
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGAAACACAACATGCGGTCTGGAACATGGCAGACAACATCAAGAAAAACAGGGTACTGCAGATCCTACCATATGACACATCCAGAGTG
E K H N M R S G N M A D N I K K N R V L Q I L P Y D T S R V
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TACATCACCCCAACTGTGGGTATGAAGAACTCAGACTACATCAATGCAGCCTTTGTTGATGGCTACAGGGAAGGATGCCTACATTGCC
Y I T P T V G M K N S D Y I N A A F V D G Y R E K D A Y I A
-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACCAAGGCCCCCTACCCACACGGTGAATGATTTCTGGAAGATGGTTTGGGAATGGAAGTCCTGCTCTATCATCATGCTGACAGAGTTG
T Q G P L P H T V N D F W K M V W E W K S C S I I M L T E L
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGGAGAGGGGCCATGAAAAGTGTCAAGTACTGCCAGGTGAGGTGGAGATGTATGGAGACATTTGTGTGGAAGCCAAGGGAGACAAA
Q E R G H E K C H K Y W P G E V E M Y G D I C V E A K G D K
-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACCTTCCAGGACTACACCGTCCGCACCTTCCACATCACTAACACAAAAGCCACAAGTTCCAGAAGCATGCTGAAGGGGAGACAGAAGGG
T F Q D Y T V R T F H I T N T K T H K F Q K H A E G E T E G
-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGCTGGAAAAGACGGCATCCTGTCGTACAGTGCAGCAGTTCCACTTCCACGGCTGGCCTGAGATCGGCATCCCAGCAAATGCTGCCGGC
L L E K T A S C R T V Q Q F H F H G W P E I G I P A N A A G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATGTTGGATCTGATTGGTCAGGTGGAGAGGCAGCAGCAACAGTCTGGAAACGGACCAATCACAGTTCAGTGCAGTTCTGGTGCAGGCCGC
M L D L I G Q V E R Q Q Q Q S G N G P I T V H C S S G A G R
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACGGGGGCATTTCATAACACTGAGCACAGTGATAGAGCGGGTGAAGGCAGAGGGCATCTGTGATGTTTTCCAAACTGTCAAATCCATGAGA
T G A F I T L S T V I E R V K A E G I C D V F Q T V K S M R
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TACCAGAGACCTCACATGGTGCAAACAGTGGAAACAGTACCAGTTCATCTATCAGGCTGTGCTGGAGTACCTGGACAGTTTCACGCTCTAT
Y Q R P H M V Q T V E Q Y Q F I Y Q A V L E Y L D S F T L Y
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCAAACCTTTAAGGACCTGCAATAGCCTGGTCTGTGGCACACAGGCTCTTT
A N F K D L Q .

```

amphioxus amPTPR5 (AB033565)

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAACAGAACACATCTGTCATCATCATGCTCACTAACCTGGTGGAGCGGGGCAGGCCGAAGTGTGACCAGTACTGGCCACAGATGGGACG
E Q N T S V I I M L T N L V E R G R R K C D Q Y W P T D G T
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGCAGTACGGGCTGATCACAGTCACCCTGAAGCACACCAAGGTGCTGGCCAACTTCACCATCAGGACATTCTCCTTACGCCACGCCAAA
E Q Y G L I T V T L K H T K V L A N F T I R T F S L R H A K
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTCAAGAAGGGAAAGGCGGAACGTACAGTGTGTCAGTTCCACTACACACAGTGGCCGGACCATGGCACCCCCGACTACGCCCTCCCTGTC
V K K G K A E R T V L Q F H Y T Q W P D H G T P D Y A L P V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTCACCTTCATCAGGAAGTCCGCTGCTGTCAAAGTGGGGGAGGATGTGGGGCCCATGGTGGTGCAGTGCAGTGTGGAGTGGGGAGGACG
L T F I R K S A A V K V G E D V G P M V V H C S A G V G R T
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGGACGTACATTGTGCTGGACACCATGATGAGGCAGGTGGAAGAGAAACATTTCGGTCAACATCAAAGGCTTCCTCAAGCACATCCGCACA
G T Y I V L D T M M R Q V E E K H S V N I K G F L K H I R T
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGAGGAACCTTGTACAGACCGAGGAGCAGTACATCTTCTGCCATGATGCCCTGCTGAAGTTTATCCAGTGTGCGGACACGGAGATC
Q R N Y L V Q T E E Q Y I F C H D A L L K F I Q C R D T E I
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCCCAGAACCAGCTGCACCACTACATGGACGAGTTGCTGAAGCCCGGGCGCTCCGGCCGGACCCTGCTGGACAAGCACTTCAAACACATG
P Q N Q L H H Y M D E L L K P G R S G R T L L D K H F K H M
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATGGCGCACCAGGCCGAGGAGTTTGAGTACTACTGTGCACAGAAGCCGTGCAACACCAAGAAGAACAGAACACAGAACTCTTACCAGTG
M A H Q A E E F E Y Y C A Q K P C N T K K N R T Q K L L P V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGTGCACCTCGTGTGTCTTACAGCAAAACCAGGGATGGAGGGTTAGATTACATCAATGCCTCTTATCTCCAGGGCTACAACAGGAGT
E C T R V C L T A K P G M E G S D Y I N A S Y L Q G Y N R S
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAGGAGTTCATCGTTACCCAGCATCCCTTGACGGACACAGTGAAGGATTTCTGGCAGATGATCTGGGACCAGAACTCCCCGGCGGTCTGTC
K E F I V T Q H P L T D T V K D F W Q M I W D Q N S P A V V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATGCTGAGCAAGCCTGACGAGAATGACAACCTGCCAGCTACTGGCCAGCAAGGACAAACCCATGGAGTGCCTGACCTTCAAGGTGTCC
M L S K P D E N D N L P S Y W P S K D K P M E C V T F K V S

```

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTGCAAGATGAGAAGGAGGAGAAAGTCGGCAACGACATCAGCGTCACTGTCCGCAGCTTCCTCCTTGAAGCCACCCAGGACGATTATGTG
L Q D E K E E K V G N D I S V T V R S F L L E A T Q D D Y V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTGGAGGTACGCCACTTCCAGTCCCCTGATTGGCCAGCGGCGTGTACGGCGATCAGCGATGTGTTCCACCTGGTGGACCTGGTGCAGGAG
L E V R H F Q S P D W P A A C T A I S D V F H L V D L V Q E
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TACGCTGCCACCAACGAGGGCCCCATCACTGTTTCATGATGTGTACGGTGGAGTGTCTGCCGCCACCTTCTGTGGGCTGACAGCTCTCAAG
H A A T N E G P I T V H D V Y G G V S A A T F C G L T A L K
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CACCAGCTAGAGGATGATGGGAGTATGGATATCTTCCAAGTGGCCAAGACCTACCACATGATGAGGCCAGGTGTCTTCTCGGAGAAGGAC
H Q L E D D G S M D I F Q L A K T Y H M M R P G V F S E K D
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGTACCTGTTTCCTATATAAGGCCATGCAGACCCAGGTGGCAGGCGTCACTGAGAACGGTTTCCGCGACATCATCAGATTGATCGCACC
Q Y L F L Y K A M Q T Q V A G V T E N G F G D I I T V D R T
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GATGAGAGTACAGAGTCGCTGGTCTGAGGTCTGCACACTGACCAGGTCTTCTGTCTCTGCTGAAAATATGTCTTCTCTGTTGTGCACTATA
D E S T E S L V .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

amphioxus amPTPR2A (AB033566)

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGCAGCGCAGCGCTACCGTCGTTATGATGACTAAACTAGAAGAGAGGACCAGGGTGAAGTGCGATCACTACTGGCCGTCCAGGGGGACC
E Q R S A T V V M M T K L E E R T R V K C D Q Y W P S R G T
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGACGTACGGGTCAGTCCAGGTCAACCTGCTGGACACCATCGAGCTGGCCACCTACACCGTCAGGACCTTCGCTCTCTTCAGGCATGGC
E T Y G S V Q V T L L D T I E L A T Y T V R T F A L F R H G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCGAGCGAGAAGCGTGAGGTGCGACAGTTCAGTTCACGGCCTGGCCGGACCACGGCGTGCCCGAACACCCGACCCCTTCCTCGCCTTC
S E K R E V R Q F Q F T A W P D H G V P E H P T P F L A F
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTCAGGAGAGTCAAAGCCTGCAACCCGCTGACGCTGGGCCCATGATCGTCCACTGCAGTGCTGGTGTGGGCAGAACCGGCTGCTTCATC
L R R V K A C N P P D A G P M I V H C S A G V G R T G C F I
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTCATCGACGCCATGTTGGAGCGCATCAAGCACGAGAACGACGGTGGACGTGTACGGCCACGTGACGCTCCTGCGCGCCAGAGGAAGTAC
V I D A M L E R I K H E K T V D V Y G H V T L L R A Q R N Y
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATGGTGCAGACCGAGGACAGTACATCTTCATCCACGACGCCATCTTCGAGGCCGTGTCCAGCGGAAACACCGAGGTCCCCGCACGGAAC
M V Q T E D Q Y I F I H D A I F E A V S S G N T E V P A R N
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTGTACGCACATATACAGAAGTTGACTCAGATTGAACCCGGAGAAACGGTGAATGGAGTTGGAATTCAAGCGGCTGGCCAATACC
L Y A H I Q K L T Q I E P G E T V T G M E L E F K R L A N T
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAAGCTCACACCTCCAGGTTCTGTCAGTGCCAACCTTCCCTGCAACAAGTTCAAGAACAGATTAGTCAACATCATGCCATATGAATCAACA
K A H T S R F V S A N L P C N K F K N R L V N I M P Y E S T
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```


-----+-----+-----+-----+-----+-----+-----+-----+-----+
 ACAAGGAAAGGATCTAGGGTAT

hantish hgPTPR2Aa (AB033569)

90

-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GAACAGCGCAGCGCTACTGTGGTCATGATGACCAAACCTGAAGAAAGATGCCGGGTGAAGTGTGATCAGTACTGGCCTAGCCGTGGTACT
 E Q R S A T V V M M T K L E E R C R V K C D Q Y W P S R G T

180

-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GAGACATATGGCCTTGTGCAAGTGACATTATTAGATACTATGGAGCTGGCAACATATGCCGTTTCGCACTTTTGCTCTGCAGAAGGCCGGT
 E T Y G L V Q V T L L D T M E L A T Y A V R T F A L Q K A G

270

-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TCAAATGAAAAACGAGAAGTTCGGCATTTCAGTTTACAGCTTGGCCTGATCATGGAGTGCCAGAGTACCCAACCTCCATTCCCTCGCATTT
 S N E K R E V R H F Q F T A W P D H G V P E Y P T P F L A F

360

-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CTACGGAGAGTTAAGGCTTGCAACCTCCTGATGCTGGACCTATGGTTGTGCACTGCAGTGCTGGTGTGGCCGCACTGGCTGCTTCATA
 L R R V K A C N P P D A G P M V V H C S A G V G R T G C F I

450

-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GTCATTGACGCCATGTTGGAACGGCTCCGGCAAGAGCACACAGTAGATGTTTATGGTCATGTTACGTGCCCTACGTGCTCAGCGCAATTAT
 V I D A M L E R L R Q E H T V D V Y G H V T C L R A Q R N Y

540

-----+-----+-----+-----+-----+-----+-----+-----+-----+
 ATGGTGACAGACAGAAGAGCAGTATGCTTTTATCCATGAGTCACTACTTGAAGCAGCAGCAGCGGGATGCACGGAGGTGCCCCGACGCTCC
 M V Q T E E Q Y A F I H E S L L E A A A A G C T E V P A R S

630

-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TTGTATGCACACCTGCAGAGGCTGACAGCACCAGGAACCTCCTGGTGAAGGCAGCTCGGGAATGGAACCTTGAGTTCAAGCGCCTTGCAAGT
 L Y A H L Q R L T A P G T P G E G S S G M E L E F K R L A S

720

-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GCCAAAGCAAGTACGCTGCGCTTTATTAGCGGAACCTCCCGTGCAACAAATTCAAGAACCGGTTGGTGAACATAATGCCGTATGAGTCC
 A K A S T L R F I S A N L P C N K F K N R L V N I M P Y E S

810

-----+-----+-----+-----+-----+-----+-----+-----+-----+
 ACGCGTGTGCTCTTCAGCCTATTCTGGGGTGTGAGGGCTCAGATTACATCAATGCCAGCTTCATTGATGGTTACAGACAGCAGGCTGCG
 T R V A L Q P I R G V E G S D Y I N A S F I D G Y R Q Q A A

900

-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TATATTGGTACTCAAGGGCCACTGGCTGAAACAACAGAAGACTTCTGGAGAATGCTTTGGGAGCATAACTCCACCATTGTTGTGTCATGCTG
 Y I G T Q G P L A E T T E D F W R M L W E H N S T I V V M L

990

-----+-----+-----+-----+-----+-----+-----+-----+-----+
 ACAAATGAGGGAGCTGGGCCAGGAAAAATGCCATCAGTATTGGCCTGCGGAACGATCTGCTCGTTATCAGTATTTTGTGTTGGATCCG
 T K L R E L G Q E K C H Q Y W P A E R S A R Y Q Y F V V D P

1080

-----+-----+-----+-----+-----+-----+-----+-----+-----+
 ATGGCCGAGTACAGCATGCCACAGTACATATTGCGGGAGTTCAAAGTCACAGATGCCAGAGATGGTCAGTCACGTACCATCCGGCAGTTC
 M A E Y S M P Q Y I L R E F K V T D A R D G Q S R T I R Q F

1170

-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CAATTCATTGACTGGCCTGAACAAGGTGTGCCAAAGTCTGGGGAGGGCTTCATCGACTTCATCGGTGAGGTGCACAAAACCTAAGGAGCAG
 Q F I D W P E Q G V P K S G E G F I D F I G Q V H K T K E Q

1260

-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TTTGGACAAGATGGGCCAATCACAGTTCAGTGCAGTGCTGGTGTGGGACGCACAGGAGTCTTCGTGACGCTGAGTATCGTTCTGGAGCGT
 F G Q D G P I T V H C S A G V G R T G V F V T L S I V L E R

1350

-----+-----+-----+-----+-----+-----+-----+-----+-----+
 ATGCGGTACGAGGGTGTGTTGACGTATTCCAGACAGTGAAGATGTTACGAACACAGCGCCCCGCCATGGTACAAACCGAGGACCAGTAT
 M R Y E G V V D V F Q T V K M L R T Q R P A M V Q T E D Q Y

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGTTCTGTTACCGTGCAGCTCTGGAATACCTGGGAAGTTTTGACCACTATGCAACATAAGGACATATGAACAATGCGGGCCATCTAGGG
Q F C Y R A A L E Y L G S F D H Y A T .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGCCACCTCAAGGTCTCTGCCAAATTTGGCGGGAAAAAATCCTCCACGTGGAATGTCGCGGGTGGGGTTTCAGCCTACAGTGTGGGCG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCATTTTGTGTTGTAAATGTTCTAGCTTATCAAAAACAAAAACAAAAAGGAACCGGGAACAAATCGCAATGATTGAGATGCACAACATGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAAGAGTAGCGCCAGGAGAGTGGTGGTGACAGCAAGAGATTAGTAGAGGAAGAAAAATAATAATATAGAAAGTAGGCAGAAACCTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGAAAGAAAGGGAAAGAGAGAAACGTTCTGTCATCGAATGTCCATGTTTTTCATCCTGTACATAACAACTGCTCAGATTCTAACATA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTCCTGAAATGGAGGGAGTTGACAGCTGATGGAGGTATGGGGTATAGAGAG

```

hagfish hgPTPR5a (AB033570)

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGCAGCAGATCGGTGTCATCATCATGATTACCAATCTTGTGGAGAAAGGTCGGAGGAAATGTGACCAGTACTGGCCGACAGAAAGCAGT
E Q Q I G V I I M I T N L V E K G R R K C D Q Y W P T E S S
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGGAATACGGTCATATCCTGGTTACTCTCAAAAGAACCAATTCTCTTGCCTGTACACCGTACGTCACTTCACGATCCGTCACATGAAG
E E Y G H I L V T L K R T N S L A C Y T V R H F T I R H M K
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTAAGAAAGGGCACACAAAAAGGAGTGGGAAAATGGCGCCAGCATGAACGAACAGTCATTCACTTCCATTACACGCAGTGGCCAGACATG
L R K G T Q K G V G K W R Q H E R T V I Q F H Y T Q W P D M
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGCACACCAGAGTATACGCTGCCTGTACTGACATTTCGTGCGCAAATCATCTGCTGCACGCACCCCTGAAATGGGCCCAGTTGTTGTGCAC
G T P E Y T L P V L T F V R K S S A A R T P E M G P V V V H
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGCGCGGGTGTGGGACGTACGGGGACGTACATCGTGGTGACAGCATGCTGCATCAGATGCGTGAACAGGGGCACGCCAATGTGCTC
C S A G V G R T G T Y I V V D S M L H Q M R E Q G T A N V L
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGGTTTCTCAAACACATTTCGCACACAACGTAACCTAGTACAGACTGAGGAGCAGTACGTGTTTCATCCACGATGTACTATTGGAAGCT
G F L K H I R T Q R N Y L V Q T E E Q Y V F I H D V L L E A
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTTTTGGCCGAAACACTGAGGTGTCTTGCCATCAGTTCCAGGCCTACTACTCTACAATTGCATCAACGCACACTGGTAATCCCTCTGAG
V L A G N T E V S C H Q F Q A Y Y S T I A S T H T G N P S E
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTGTTTCTAGCAACAGCACTTGCCTTGAGAGACAGTTCATGCTGGTGCTTCAGATGTCATTTCGACATTGCAATGTTTCGGTGCTCAA
L F P S N S T C L E R Q F M L V L Q M S F R H S E C F G A Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAGCACTGCAACAAGGAAAAGAACCCTAACTCTTCAGTTGTACCTGTGGAGCATGCTCGGGTGGGCCTTTTCATCTCTCACAGGTGTTGAA
K H C N K E K N R N S S V V P V E H A R V G L S S L T G V E
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGCTCAGATTATATCAATGCCTCATACATTATGGGATATCGTTGTAGCAATGAGTTTCATAATAACACAGCACCCCTTTACCACACACTACG
G S D Y I N A S Y I M G Y R C S N E F I I T Q H P L P H T T

```

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+
UGTGACTTCTGGGCGATGATCTGGGACCACAACCTCTGAGACTGTCGTGATGCTGCCTGGGAACCAAGGCCTGGCAGAGGATGAGTTTGAA
R D F W R M I W D H N S E T V V M L P G N Q G L A E D E F E
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TATGGCCAGGAAGAGAAGAAGCCATGACGTGTGAGGATTTTACTGTCACTAATTAATGAAGACCGGATGTGCATGTCCAATGAAGAC
Y W P G R E E A M T C E D F T V T L I N E D R M C M S N E D
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGCTCGTTATCCATGACTTCGTTCTCGAGGCCACACAGGATGATTATGTACTTGAAGTGCCTCACTTCCAATGTCCTCCATGGCCTAAT
Q L V I H D F V L E A T Q D D Y V L E V R H F Q C P P W P N
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCTGATGGGCCACTCAGCCGTGCATTCGAGCTCGTCAATTTGGTTCGAGAGGCTTCGGCAGCCACGATGGGCCAACGGTTGTGCATGAC
A D G P L S R A F E L V N L V R E A S A A H D G P T V V H D
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGTTTCGGGGGTGTAACAGCTGGCACATTTTGTGCATTGAGCACTCTCTACCAGCAACTGGAAGACCAGGCCATTGTGGATGTATTTTCAT
E F G G V T A G T F C A L S T L Y Q Q L E D Q A I V D V F H
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTGGCCAAGATGATCAACCTCATGCGTCCTGGTGTATTTACAGAGCTGGAACACTATCATTTTTTGTACCGCATGCTGCTCAGCCTCTTT
V A K M I N L M R P G V F T E L E H Y H F L Y R M L L S L F
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGCCTACAAGAAGCTGCTGGGTCTGCTGATTGCAACGGATTGCCAATTGCTAACAACGATGAAGCGGAAATGCAAAGCACAGAGTCTCTC
G L Q E A A G S A D C N G L P I A N N D E A E M Q S T E S L
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTGTAGTCGTGGCAATGAGGACACTCTTCAGACTCGTCCGTTGCATGGACAGTACCAACCGATACTCTGTCTTAATGAAGCAGCTGAAGT
V
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GATGATACTTGGGAGTTTTAATTTGCCAAGAAGACTGCCAAAACAATGCTGTTATCTTTAAACTGCAAACCTTTGTTTACAAACATGAAAA
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCACTGTACTCCAGCTGCGGTCCTTTTCATCTAACACCATGCTCGTGTTGCATCACCAGAGCTGAAGGGTACAAAACACTGCAAGGATTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGATGGTAGGCTAAAAAATGAAAAAATTGATATTGTTTGTCTTGCCTTTCACAAACCTTACAAAGTACCACAGACGGTACAGGCTGA
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGCGTAATGATTACAGAAATGTTTCATAACGCTGACGGTAATAAGATGATCTTG

```

hagfish hgPTPR2Ab (AB033571)

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGCAGAGAAGTGCCACAATTGTTACGATGACCCGACTTGAAGAGAGAGGGCGGGTGAATGTGACCAGTATTGGCCCTCTCGTGGTGTG
E Q R S A T I V T M T R L E E R G R V K C D Q Y W P S R G V
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGAACTACGGGCAGATGCAGGTTTCCTTGATCGACACAATGGAGCTGGCCACATATTCATCCGCTCCTTCGCTCTGTGCAAGGTCGGA
E N Y G Q M Q V S L I D T M E L A T Y S I R S F A L C K V G
-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGCACAGAGCAGCGGGAATTGAAGCAGTTTCAGTTCACGTCATGGCCAGACCACGGCGTTCAGAACACCCAACACCATTCTATCATTT
S T E Q R E L K Q F Q F T S W P D H G V P E H P T P F L S F
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTACGCAGAGTGCGCCAGGCCGATCCACTAGATAGTGGGCCAATGGTTGTGCACTGCAGTGCTGGAGTGGGACGCACCGGTTGCTTCATC
L R R V R Q A D P L D S G P M V V H C S A G V G R T G C F I

```

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTTTTGGAGCCCATGCTGGAGCGCATAAGGCATGAAAATACCGTGGACGTCTATGGACTTGTTACGTGTTTACGGACACAGCGAAACTAC
V L D A M L E R I R H E N T V D V Y G L V T C L R T Q R N Y
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACGGTGCAAACGGAGGAACAGTATGCATTTGTTTCACGAGGCACTGCGCGAGGCCGTGATGTGCGGTTGTACAGAGGTGCCGGCTCGTGCC
T V Q T E E Q Y A F V H E A L R E A V M C G C T E V P A R A
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTGCACACTGCTCTGCAGCGGCTGCAGCAAGTGGCACCCGAGAAAACCGCCACTGGGATGGAACCTCGAGTTCAAGCGCCTGTTTAACTCC
L H T A L Q R L Q Q V A P G E T A T G M E L E F K R L F N S
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAGCCACAGACTTCTCGGTTTCATCACAGCAAACCTTCCCTGCAACAAGTTTAAAGAATCGGCTCGTCAACGTCTCCCGTATGAGTCCACA
K P Q T S R F I T A N L P C N K F K N R L V N V L P Y E S T
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGAGTCTGCCTTCTTCCCATCCGTGGTGTTGAAGGCTCCGATTACATCAATGCGAGCTTCATTGATGGTTACCGGCAACAAAAGGCATAC
R V C L L P I R G V E G S D Y I N A S F I D G Y R Q Q K A Y
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATTGCCACACAAGGTCTCTGGCTGAAACCACAGAAGACTTCTGGAGGATGCTCTGGGAGCAGAATTCACCATTGTGGTGATGCTAACC
I A T Q G P L A E T T E D F W R M L W E Q N S T I V V M L T
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAGCTTCGCGAAATGGGCCGGGAAAAGTGTTACCAGTACTGGCCTGCAGAGCGCTCGGCACGATACCAGTACTTTGTGGTGGACCCGGTG
K L R E M G R E K C Y Q Y W P A E R S A R Y Q Y F V V D P V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCAGAATACAATATGCCACAGTACATCCTACGTGAGTTCAAGGTCAGTATGCCAGGGATGGACAATCACGGACTGTGCGTCAGTTCCAG
A E Y N M P Q Y I L R E F K V T D A R D G Q S R T V R Q F Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTTATAGACTGGCCGGAGCAGGGCACACCTCGTACAGGAGAAGGTTTCATTGACTTTATTTGGTCAGGTGCATAAGACCAAGGAACAGTTT
F I D W P E Q G T P R T G E G F I D F I G Q V H K T K E Q F
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGCCAGGATGGGCCGATAGCAGTGCAGTGCAGCACAGGTGTGGGTCCGACAGGTGTGTTTGTGGCGCTCAGCGTGGTGCTTGAGCGGATG
G Q D G P I A V H C S T G V G R T G V F V A L S V V L E R M
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGATATGAGGGCATGTTTCGACCTGTTTCAGACTGTGAAGATGATACGCACTCAGCGACCTGCAGTCTGTCAGATAGAGGAACAGTATCAA
R Y E G M F D L F Q T V K M I R T Q R P A L V Q I E E Q Y Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTCTGCTACCGGGCAGCACTCGAGTACTTGGGCAGCTTTGACCACTATGCTGTATGATGGACCAGCCAACCTGTGCGACAGTCTCTATCG
C Y R A A L E Y L G S F D H Y A V .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATCTTGTATACATCGGAATCTGTTGGTACAGCCTTGCAAGTTCAGACCTTTTGTGTGTATGAAGTCCGGCTGGGACGTCTCTGAGCGTGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAGTGGGCGATTTTCTGCTCGTACCTAAGCCACGTACCACATGCTGTGTTGGAGCTTACCCTGATTTCTCACATGGGACTAGAGTCGGCT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCGTGTCGACCTCATTCAACTATCAATTAAGGAGACTATTTGAATATCAACATACCAGCATCTGTTCCATAATGGCAGACCATCAGAAA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACAGCCAGGTCATTACTACCATTCACTTACTTTTACCATGGGCATTTGTTCAATGAACTTATGTATTTAATATGTACCATAAAATCGTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTTATGATTATTTAAATTTGTGGCCAAATGTTAAAATATGGTGTGCTGTAGGTCATGGAGAGTGTGCAGGCAACAATG

```


hagflsn ngPTPR4 (AB033572)

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGTGAGATCCACGGTGACATCAAAGTGACTCTGGTGGAGACAGAGATGCTGGCTGAATACATGATCCGCACATTGAGCCTGCATAAGGCA
S E I H G D I K V T L V E T E M L A E Y M I R T F S L H K A
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 180
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGATCCGATGAAACGAGGCAACTCAGGCAGTTCCTTCACTGGCTGGCCTGACCATGGCGTTCCCTATCATGCTACAGGCCTGTTAGCG
G S D E T R Q L R Q F H F T G W P D H G V P Y H A T G L L A
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 270
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTCATCCGCAGGGTCAAGTCTACGAATGCTCCTCATTCGGGACCCGTGGTGGTCCATTGCAGTGTGGTGCTGGGAGGACGGGCTGTTTC
F I R R V K S T N A P H S G P V V V H C S A G A G R T G C F
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 360
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATCGTCATCGATATCATGATGGACATGGCTGGGCGGGAAGGTCTCCTCGACATCTACAACCTGTGTGTGTGAACTGCGCAGCCGACGTATC
I V I D I M M D M A G R E G L L D I Y N C V C E L R S R R I
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 450
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AACATGGTCCAGACAGAGGAGCAGTACGTGTTTATCCACGACGCGATGTTGGAGGCTTGCTCTGTGGTGACACCGCCATCTCTGTGCAT
N M V Q T E E Q Y V F I H D A M L E A C L C G D T A I S V H
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 540
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACCTTCCGTGCTGCATACCCACCATTTGTATGCACGGACCCGAGAGTGGCTCTTCACAAGTAAAAGATGAGTTTCAGGTGTGTTTGCAG
T F R A A Y P T I V C T D P Q S G S S Q V K D E F Q V C L Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 630
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CATTTGCTTTGCATCTTGCTCATATATCAGAAGTTTTTCATCCTTTGCATTGAGTGAACATTGCGATTGCTGGAGGTGTGCATGCTGTGCA
H L L C I L L I Y Q K F S S F A L S E H S H C W R C A C C A
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 720
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTCTAATAAGTAGTGTAGCAGAAATTGTATGATTATTCAGGGACAGGTATCCATGGTAATCCATGCCCCCTCCATGAAACCTTGGTATA
F .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 810
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCTAATGTGCCGTATATTGCTCAGAATTGAAATGAATCATGAAAAAAAAAATTTCAACATGAACTGTGACAAACCGTCAGCCAACTTAT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 900
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGCTCCTGGGATCAGGGAATGTCCTTGTCTGTGCCTGGTCTCTCTTTGAGTCTCTCTACTCTTTCTGCTGCTATTCAAATTGAT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 990
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GACTGGTCACTTATCAGGTGCAATGACTTGTACCTGATAATGACTCCACATGAAGGTGCATTGATTAATGTGACTGTTGTGCTCATGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1080
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTCTCCTTGGTCTTTTGTTCAGCATAATAAGATTTTACATTTGAAAAAAAAA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1170

hagfish hgPTPR2Ac (AB033575)

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAACAGCATTGCAGTACCATCATCATGATGACCAGGCTTGAGGAGAGAGGAAGGGTGAAATGTGACCGGTACTGGCCTCTTCGTGGTGTG
E Q H C S T I I M M T R L E E R G R V K C D R Y W P L R G V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 90
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGACATATGGCTTGGTATCTGTGACTCTGTTGGATACACAGAGGTGGCGAGTTATTGTCTTCGATCCCTCACCTTGTGTAAGGTGGGA
E T Y G L V S V T L L D T T E V A S Y C L R S L T L C K V G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 180
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGCAGGGAAAGCGGGAAGTGAAGCATTTCCAATTCTTGGCCTGGCCAGATCATGGAGTCCCAACCACCCAGCCCCTATTCTTGCCTTC
G R E R R E V K H F Q F L A W P D H G V P N H P A P I L A F
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 270
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTGAAACGTGTTTGGGCATGCACCTCTACAGAGACTGGGCCTCTGGTGGTTCACTGCAGTGTGGGATTGGTGAACAGGTTGCATTATT
L K R V W A C T S T E T G P L V V H C S A G I G R T G C I I
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 360

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCACTGGATACAGCTCTGTCATCGCCTGCGTGCGGAGCGCTGTGTGGACGTGTATGGCCATGTGACCTGTCTGCGGAGCTCACGCCCTGGC
A L D T A L H R L R A E R C V D V Y G H V T C L R S S R P G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACTGTGCAAAGCCAAGAGCTCTATGGGTTTGTGTACTCTGCGGTACGTGAGGCCATTGGTGAGACCACAGCTGGAGCAGCCACAGAGATC
T V Q S Q E L Y G F V Y S A V R E A I G E T T A G A A T E I
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTGGTGCGCCTCCTTCACACCAGGCTCCAGCGGCTGCAGCAGCGGGTGCCTGGGGAAAACGTTACAGGACTCGAGCTGGAATTC AAGCGT
L V R L L H T R L Q R L Q Q R V P G E N V T G L E L E F K R
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTAGGAGTGAGAGTGCTCTACTTCATGCCCACCAGTCGTGCTCGCGCTGCCCTGCAATGCGCAAAAGAACCGTCCGGGCGCGCTGCTGCCA
L G V E C S T S C P P V V L A L P C N A Q K N R P G A V L P
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TACGAGCACAACCGGGTTCGTCTGCAGCATGTACGTGTCCAACACTCCACAGACTACATCAATGCCAGTCTGGTCGATGGTTACAGGCAT
Y E H N R V R L Q H V R V Q H S T D Y I N A S L V D G Y R H
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGGCGAGCATACATCGTCACTCAGGGCCCTCTTGCTGGAACAACGGAAGATTTCTGGCGCCTGGTGTGGGAACAAAACCTCGCCATAATC
R R A Y I V T Q G P L A G T T E D F W R L V W E Q N S P I I
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTTATGCTGACCAAACCTAAGAGAGATGGGCAGGAAAAGTGCCACAGGTACTGGCCAGCAGAGCGTGCGGCACGTTACCATTTCTGGTG
V M L T K L R E M G R E K C H R Y W P A E R A A R Y H F L V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTTAGCCCCATTGCTGAGTACAACATGCCACAGTACATCCTGAGAGAATTCAAGGTGACAGATGCAAGGGATGGCCAATCACGGACGGTA
V S P I A E Y N M P Q Y I L R E F K V T D A R D G Q S R T V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGACAGTTCCAGTTCAGCGACTGGCCAGAACAGGGTGTACCACGTATGGCTGAGCCGTTCTCGAGTTCCTTGGTCAGGTGCATAAAACG
R Q F Q F S D W P E Q G V P R M A E P F L E F L G Q V H K T
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAAGACCAGTTTGGGCAGGATGGGCCTATCACAGTGCATTGCAGTGCTGGAGCTGGTCGTAGCGGTGTATTTCGTTACCATTAGTGTGCTT
K D Q F G Q D G P I T V H C S A G A G R S G V F V T I S V L
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTGGAGCGCATGAGGAGTGAAGGTGTGCTTGATGTTTTCTTAAGTGTCAAGATGCTTCGTCTGCAACGGCAGGGGATGGTCCAAACTCAG
L E R M R S E G V L D V F L T V K M L R L Q R Q G M V Q T Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAACAGTACGAGTTTGTATCGAGCGGCACTGGAGTACCTGGGAAACTTTGACCACTTTGGAAACTGAGCCGAAACAGCTCAACATCTA
Q Y E F C Y R A A L E Y L G N F D H F G N .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGTGACTTAAAACAACATGCCAAAAGAAGACAACAGCAATGGACCCAAAACAACAATGGGATACTATTTGTAAAAGGATTACAAAGAAAA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCAATGGTGTCAACCACAACAGTACAATAATATACCGACTTTTAATGATG

```

hagfish hgPTPN6a (AB033576)

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTTTACCAAGAGAACTCAAGAGTTATTTTCATGACCACAAAGGAGGTGGAAAGGGGTCGGAGCAAGTGTGTTTCGCTACTGGCCTGAAGAG
V Y Q E N S R V I F M T T K E V E R G R S K C V R Y W P E E
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGCTGTAGCAAAGACTACGGCTATCTTCGAGTGTAACAAGTACACGAATCAGAATCCTACGACTACGCTCTGCGCCAGCTGGAGCTCACA
G C S K D Y G Y L R V Y N V H E S E S Y D Y A L R Q L E L T

```

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CATGTGGAGCATGCTGAACCTACGGATAGTGTGGCAGTACCACCTTTAAGGCATGGCCAGATCACGGAGTACCCAACGAGCCAGGTGGTGTG
H V E H A E L R I V W Q Y H F K A W P D H G V P N E P G G V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTCAACTTCCTGGATGAGGTTAACCGGCGCCAGGATTCCATCCCTGATGCTGGGCCGCTCGTGGTTCATTGCAGTGCCGGGATTGGTTCGA
L N F L D E V N R R Q D S I P D A G P L V V H C S A G I G R
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACAGGAACCTTCATTGTCATCGACATTCTCATCGACATCATTGCCACAAAGGCCTCGATTGTGATATCGATATCCCAAAGACGATCCAG
T G T F I V I D I L I D I I R H K G L D C D I D I P K T I Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATGGTGAGGGCACAGCGCTCGGGGATGGTGCAGACTGAGGCACAGTTCAAGTTTGTGTACATGGCTGTGCAGCAATACATTGAGACAATG
M V R A Q R S G M V Q T E A Q F K F V Y M A V Q Q Y I E T M
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGCAGCGACTCGCTGAAGAGCAGAAGAGTAAAGTGAAAGGCAGAGAATACACGAATATCCGTTACTCTGCAGCTGACCTTGGGTGTCCG
Q Q R L A E E Q K S K V K G R E Y T N I R Y S A A D L G C P
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAAGGGTCTTTTCCACCACCGACGCCCATCCGTCTCTCTGCTCTGTGAGTCACCTTTAAACATTGCACGATATACTTTCCTCTGGTTATT
E G S F P P P T P I R L S A L .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTCCCGTCGCTCCATTACCTAGCCCTTTGTTCACTTTGATGGCAAACACTTGTGGTTCGCTTGTGTTTGTTTATTCATGATTTCCAAGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAAAAAGTAAGACCTTTCTACGGTGGGACACCGCTCAATCATGCTGATGTGTTCCGCAGTCTAATAGAACTAACGGGAATAACATATTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTGTTTCGCACAATCTTGCCCGAGATACGGTACATATTGTTTGTTCATAATAACACTTTGATTTTTAGTTTTATACGATGGTATGTATA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACCGTATATAAATGAACATGAATAATGATTTTCGACAAAAAATAGAGTGAGGGCAAAAAAAAAAAAAA

```

hagfish hgPTPR5b (AB033577)

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTGTGGGAACACGACGTCAGAATCATCATCATGATAACCAATCTTGTGGAGAAAGGAAGGAGGAAATGTGACCCTTACTGGCCGATGGAG
V W E H D V R I I I M I T N L V E K G R R K C D P Y W P M E
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGCGGTGAGGAATATGGGACTGTGGTCGTGACGCTACGTGCCACTCGTACACTGGCCTGCTACACCGTGCGGCGTTTCACGTTACGTCTC
G G E E Y G T V V V T L R A T R T L A C Y T V R R F T L R L
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACCAAGCCTCGCCGGGCAAATCAACGAGGAGTGGAGCGAACGGTCCTCCAATTTCAATACACCTCCTGGCCAGACATGGGCACACCCGAG
T K P R R A N Q R G V E R T V L Q F Q Y T S W P D M G T P E
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CATGCTCTGCCTCTGCTTGCCCTTCGTACGGACGACGGTGCAAGCAGCAGCACAAAGCCGAGCGCAGGCCTCGGTGCAGGCCACGCTGCAA
H A L P L L A F V R T T V Q A A A Q A A A Q A S V Q A T L Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCCTCATCAAAGTCACAGACTCGAAGCAAGAAGAAGCCACGGGTCGTGTGGGTCCAGGGCCTGTGGTGGTGCATTGCAGTGCGGGTGTG
A S S K S Q T R S K K K P R G R V G P G P V V V H C S A G V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGGCGTACAGGCACGTACATAGTCCTCGCATCGATGCTGCAGCAGATGAAGGAGAAGGGAACCATGGTATCAAGGGCTTCTTGAAGCAC
G R T G T Y I V L A S M L Q Q M K E K G T I G I K G F L K H

```


-----+-----+-----+-----+-----+-----+-----+-----+
ATCTACCAGGAAGGCTCTCGAGTTATCGTCATGACCACGAGGGAGGTGCAACGTGGCAAGGTGAAGTGTGTGCGGTATTGGCCGGAACCTT
I Y Q E G S R V I V M T T R E V E R G K V K C V R Y W P E L 90

-----+-----+-----+-----+-----+-----+-----+-----+
GGAATCGAGAAGGAGTTTCGGAATGCTGAGGGTGGGCCTTGTTGATGAAACGGCAAAGCAAGAGTACACTCTCCGGGAACCTTCCCTGCAG
G I E K E F G M L R V G L V D E T A K Q E Y T L R E L S L Q 180

-----+-----+-----+-----+-----+-----+-----+-----+
CACATCGGCTTTCCCACGGCGGTACGCACCATTTGGCAGTACCAGTACAAGGCTTGGCCCCGATCACGGTGTGCCAGGTGACCTTGGAGAG
H I G F P T A V R T I W Q Y Q Y K A W P D H G V P G D L G E 270

-----+-----+-----+-----+-----+-----+-----+-----+
GTGATCGAGTTTTTGGATGAGGTGGCGAAACGGCAGGCATATCTACCCGAGGCTGGACCCGTCATCGTGCAGTGCAGTGCCGGAATTGGT
V I E F L D E V A K R Q A Y L P E A G P V I V H C S A G I G 360

-----+-----+-----+-----+-----+-----+-----+-----+
CGAACAGGAACATTGATCGTCATAGATATCCTGACCGACATGATCCTGCACAAAGGCCCAAACCTGCAGCATTGACATCCCATGGGTAGTG
R T G T L I V I D I L T D M I L H K G P N C S I D I P W V V 450

-----+-----+-----+-----+-----+-----+-----+-----+
CAGTGGGTGAGGTCCCAGCGCTCAGGCCTGGTACAGACGGAGAGTCAGTTTCGTTTCATATACGAGGTGTAATGCACTTCATAGAGAAT
Q W V R S Q R S G L V Q T E S Q F R F I Y E V V M H F I E N 540

-----+-----+-----+-----+-----+-----+-----+-----+
CTCAAGAAAAGAGGGCAAACCAACAAGAAATGAACTTCAAACCTGGCGAGTACTCAAACATCGATGCTGCAACGCAAGACATGAGACGTCCA
K K E G K P T R N E L Q T G E Y S N I D A A T Q D M R R P 630

-----+-----+-----+-----+-----+-----+-----+-----+
ACTTCTGGCAGCCGTTTATCTCTCAAAAAAACAAGGAGAGAGAATCAGGGGACACTGTATGAAAACGTCAACCGCAAACACAAACCA
T S G S R L S L K K N K G E K N Q G T L Y E N V N R K H K P 720

-----+-----+-----+-----+-----+-----+-----+-----+
TCGGGAGTTTAAACAGCGCGAGCCACCGCTTGCATTGCGGGATCCAATCGCTCCTCATGCGAGCTCTTTCACAGTGACGAGGGGCTCGTCG
S G V 810

-----+-----+-----+-----+-----+-----+-----+-----+
TCGTTACCGAGTCTGCGCAAACACGTGCCGTCTGCTTTCCACAACGCCCAGCATGGCTCCACATTTAACAATTTGGTTTGGACGCC 900

-----+-----+-----+-----+-----+-----+-----+-----+
ATTACAGACGGGGTACCGTGACAAACCACATCCGTGTTTCGTGGAACAATAAATAAAAAAGAAGCTTTTATTGAAAAAAAA 990

hagtlsh hgPTPN3 (AB033579)

90

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAACAACAGTACACGGGCAGTGGTTATGCTGAACCGTGTGCTTGAAAAAGGCTTGGTGAAGTGTGAACAGTATTGGCCGACAGTCGATGAG
E Q Q S R A V V M L N R V L E K G L V K C E Q Y W P T V D E
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
180
CCAGAGATGGTGGTTTGACACCGGCTACCGACTCACACTGGCTTCATCAATACCAAAAGAGTACTTCACCATCAGCCGTCTAATGCTGGAG
P E M V F D T G Y R L T L A S S I P K E Y F T I S R L M L E
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
270
TGCTTATCGACGGCGGAGAGACGGGAGGTTTGCACCTCCACTACACGTCGTGGCCTGACTTCGGTGTGCCCCGACTCACCAGCGGCATTC
C L S T A E R R E V L H F H Y T S W P D F G V P D S P A A F
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
360
CTGGCATTCCTGATGCGTGTGCGGCGCTCAGGTGTACTCGATCCCTCTTTTGGTCCAATTGTTGTGCACTGCAGTGCCGGCATCGGGCGT
L A F L M R V R R S G V L D P S F G P I V V H C S A G I G R
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
450
TCGGGCACGTTCTGCCTCGTAGATTCTGCCTTGCTCTGCATGCAGCCCGGGGCGAGCCAGACGTGTACCAGGTGCTGCTGGACCTGCGA
S G T F C L V D S C L A L H A A R G E P D V Y Q V L L D L R
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
540
AGGTTTCGTATGGGTTTGATCCAAACCCCGGCTCAGCTGCGGTTCTCATTCCTTGCCATCATCCAGGGCATGCATCTTTTGAGTGCAGAT
R F R M G L I Q T P A Q L R F S F L A I I Q G M H L L S A D
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
630
GCCTGTGGCCAGGAGGGCTATTTGGATGAAGAAGATGAAAATATGTCACCTGAGGAATGGCCATACTTTGCCCTTTTCCTTCACCACCT
A C G Q E G Y L D E E D E N M S P E E W P Y F A P F P S P P
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
720
TCCACCCCTCCTCCCCACCCCCACGGCCTCCAAAACGCTCGACAGAGAGACCCAATGGAGGCCCTGCAGCTGAGAAGTCAATGGTCGTT
S T P P P P P P R P P K R S T E R P N G G P A A E K S M V V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
810
GAAGATCTAACACAACTGAATTATTTCCAGATAACAGAGATTATGAACTGTCCTCACCAACTCGTTATGAATTCAAAGTGGACGATGGC
E D L T Q T E L F P D N R D Y E L S S P T R Y E F K L D D G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
900
AGGTGTGTCGATGAGGCTTTTCGGAGAGCGTTGCAGGGAGGTTCGAGCAGACTGAGCAGACGCTTAAAGTTGATCGGGAAGAGAAAACCAAT
R C V D E A F G E R C R E V E Q S E Q T L K V D R E E K T N
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
990
GGGGAACAAAAGATCAGTGAGGATGAGGCCTGGAGCTGGGAGGCGAGCACACTGCCCAAAACAAAGACTAGATCAGACAGCAGCATGGGC
G E Q K I S E D E A W S W E A S T L P K T K T R S D S S M G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
1080
ATTGACGCAGGAAATGGCAGCTGCACACTGGCTGACAGGGGCCAAATGGGCAGTTGGGGTTCGAGAGAGGTTGGCAAGACCCACTGACG
I D A G N G S C T L A D R G Q I G Q L G F E R G W Q D P L T
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
1170
CAAATTGTCCTTATTGTTTTTGCTTTACTGCTTTACGCGGCCTACCGCCACTGCATATAAGGTTGGCCTTGTGTAACATGCACCCCGATA
Q I V L I V F A L L L Y A A Y R H C I
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
1260
CCTTTTACTACCTGGCCAAACTCCAGACAGGTGTTGTCACCCACCTATTCCAGAGCGAGCCAGTTCTGTAGCAATAGTACACATCACCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
1350
GCTAC

ray ryPTPN6a (AB033580)

90

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGGAGAACACCAGGATCATTGTGATGACCACTCGCGAGGTGGAGAAGGGACGGACAAAATGTGAACCTTACTGGCCAGAGGAACAGGGG
Q E N T R I I V M T T R E V E K G R T K C E P Y W P E E Q G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
180


```

-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAAGCACATACCTCACGATTTATCAGTGCUAATCTCCCTTGTAATAAATTTAAAAACCGTCTTGTGAACATCATGCCATATGAATCCACT
K A H T S R F I S A N L P C N K F K N R L V N I M P Y E S T
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGAGTCTGCTTGCAACCAATCCGAGGAGTTGAGGGCTCTGATTACATAAATGCCAGCTTCATTGATGGATACAGGCAGCAAAAGGCCTAC
R V C L Q P I R G V E G S D Y I N A S F I D G Y R Q Q K A Y
-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATTGCAACACAGGGACCTCTGGCAGAGACCACGGAAGATTTCTGGCGTATGCTATGGGAGCACAAATCCACAATTGTTGTAATGCTCACA
I A T Q G P L A E T T E D F W R M L W E H N S T I V V M L T
-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAATTACGAGAAATGGGTCGAGAGAAGTGCCACCAGTACTGGCCAGCAGAACGTTCCGGCCCCGGTATCAATACTTTGTGGTAGATCCAATG
K L R E M G R E K C H Q Y W P A E R S A R Y Q Y F V V D P M
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCAGAATATAGCATGCCACAGTATATCCTGCGAGAGTTCAAAGTAACAGATGCCAGGGATGGACAATCTCGAACATCCGACAGTTTCAT
A E Y S M P Q Y I L R E F K V T D A R D G Q S R T I R Q F H
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTCATTGACTGGCCTGAACAAGGAGTGCCAAAAACTGGAGAAGGATTCATTGACTTTATTGGACAAGTACACAAGACAAAAGAACAATTT
F I D W P E Q G V P K T G E G F I D F I G Q V H K T K E Q F
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGACAAGATGGACCCATTACAGTGCAGTGCAGTGTGGTGTGGGACGAAGTGGAGTATTCATTACACTAAGCATTGTTCTGGAGCGAATG
G Q D G P I T V H C S A G V G R T G V F I T L S I V L E R M
-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGATATGAGGGCGTTGTGGACATGTTTCAAAGTGTGAAGACATTACGAACACAGCGGCCTGCCATGGTTTCAGACAGAGGACCAGTACCAG
R Y E G V V D M F Q T V K T L R T Q R P A M V Q T E D Q Y Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTATGTTACCGAGCCGCACTGGAGTACCTTGAAGCTTTGATCACTATGCAACGTAACCTCTAGCACCCAATCAGAACATATGATTTTCA
L C Y R A A L E Y L G S F D H Y A T .
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGACCCAGAAATGATTTTTTGGAGCCATATCATATTAAATCATTGTTTCAG

```

ray ryPTPN3 (AB033582)

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGCAAAAAAGTAGAGCAATTGTGATGCTCAATCGTGTCAATTGAAAAGGGATCAGTAAATGTGAAAAGTATTGGCCAAGCACGAAAGAT
E Q K S R A I V M L N R V I E K G S V K C E K Y W P S T K D
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAAACAATGTCATATGAAGATACTGGGTTCAATCTAAAATTGTTATGTGAAGATGTAAATCATACTACACTGTGCGGCGGTTGGAATTG
E T M S Y E D T G F N L K L L C E D V K S Y Y T V R R L E L
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGAGAACCTCTACTGGAGAAACCAGAGAAATCTACCACTTCCATTATACCACTTGGCCTGACTTTGGAGTACCTGAATCCCCTGCTTCA
Q R T S T G E T R E I Y H F H Y T T W P D F G V P E S P A S
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTCCTTAATTTTTTGTTCAAAGTGAGGGAGTCTGGCTGCTTGAAGTCAAGTGTGGACCTACTATTGTTCACTGCAGTGCAGGCATTGGA
F L N F L F K V R E S G C L N S D C G P T I V H C S A G I G
-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGATCTGGTACTTTCTCTTTGGTTGATACATGCCTTGTTTTGATGGAAAAACGCAAGATCCATCTTCTGTGGACATCAGAGAAGTACTG
R S G T F S L V D T C L V L M E K R K D P S S V D I R E V L
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTAGAAATGAGGAAATACAGGATGGGACTCATACAGACTCCAGATCAGTTGCGATTTTCATACATGGCAGTTATAGAGGGAGCGAAATAT
L E M R K Y R M G L I Q T P D Q L R F S Y M A V I E G A K Y

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 ATCATGGGAGACTCATCAGTGCAGAAAAGATGGAAAGAGCTTTCCAACGAAGACCAGGAGCCTATACCAGGCAATGCTTCTCCACCACCC
 I M G D S S V Q K R W K E L S N E D Q E P I P G N A S P P P
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 720
 TAGCCTCCAAAACGAGCCCTTCCACGACAGAATGAGAATAGGGCGCCTCATTTGTCTAGTAGGCGAAGCAACTGATGCAGCAAACAACATT
 Q P P K R A L P R Q N E N R A P H L S V G E A T D A A N N I
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 810
 TCCACAGAAGAAAATTCTGTACAAGAAAAAGAAATCAAAGCAAAGTATGTGAATAATATTCGTAAGCGACATCGTGAGGATAGGATAGCA
 S T E E N S V Q E K E I K A K Y V N N I R K R H R E D R I A
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 900
 ACTACAGCGCAGAAAAGTACAGCGAATGAAACAGAAGCTCAATGAAACCGAGGTGAAACGAAAAGGTGGTTATATTTGGAACCTATTCTC
 S T A Q K V Q R M K Q K L N E T E V K R K R W L Y W K P I L
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 990
 ATTAACATAGCTGCAGTTACTACTCTTGCAATTGGAGCTATTGTTTGCTGCTCTATGTATTTTCAGTGATCAATTTGTAACCTTCATTTAA
 I N I A A V T T L A I G A I V C C S M Y F Q .
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1080
 TAGCACATGGTAGCATATTGCATCTGAGATTTAGATGCTGAGGAAAGATATCTCCCCTGCACAAGAACTTTGTTTTTCAGGATTGACAG
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1170
 AGAAAAATACAGGAGGGTGGAGCTGTTTGTTTACTCAAAATTTGCAATATAGAATCTCAGGTTCTTGTACTGTCTGATGTTTGAATGA
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1260
 ACAAACACCAGACACAAAAATT

-----+-----+-----+-----+-----+-----+-----+-----+
GAACAGAAATCTGCCACCATTGTTATGTTAACAAATGTAAAAGAGAAAAAGGAGGAAAAGTGTTACAAGTACTGGCCTGATCAGGGATTA
E Q K S A T I V M L T N V K E K K E E K C Y K Y W P D Q G L 90

-----+-----+-----+-----+-----+-----+-----+-----+
TGGACGTATGGGGCAATTCGGGTCTCAGTGGAAGATGTCACTGTGCTTGTGGATTATACCATACGCAAGTTCACCATTGAGAATCAGACA
W T Y G A I R V S V E D V T V L V D Y T I R K F T I Q N Q T 180

-----+-----+-----+-----+-----+-----+-----+-----+
AQTGTTAACAAAGCTCCAAGACTAGTGACACAGCTTCACTTCACCAGCTGGCCTGACTTTGGTGTGCCATTTGCACCCATCGGAATGCTC
V N K A P R L V T Q L H F T S W P D F G V P F A P I G M L 270

-----+-----+-----+-----+-----+-----+-----+-----+
AAATTTCTGAAGAAGGTTAAGGCATTGAATCCTTCCTATTCTGGGACCTATTGTGGTCCACTGCAGTGCTGGAGTGGGCAGAACTGGAACC
K F L K K V K A L N P S Y S G P I V V H C S A G V G R T G T 360

-----+-----+-----+-----+-----+-----+-----+-----+
TTCATTGTTATTGATGCTATGATTGATATGATGGATGCAGAGCAGAAGGTGGATGTCTTTGGCTTTGTGGCTCGAATTCGCAATCAGCGG
F I V I D A M I D M M D A E Q K V D V F G F V A R I R N Q R 450

-----+-----+-----+-----+-----+-----+-----+-----+
CCACAGCTAGTTCAAACCTGACATGCAGTATTCCTTCATATATCAAGCCCTTCTTGAACATTATTTGTACGGTGACACAGAACTGGATGTG
P Q L V Q T D M Q Y S F I Y Q A L L E H Y L Y G D T E L D V 540

-----+-----+-----+-----+-----+-----+-----+-----+
AGCTCTTTAGAGAAGCACATACAGAAGCTCCGAAATAAACAGGGAAATTTGACATCACTGGCCTAGAAGAAGAATTTAATAAATTAAC
S S L E K H I Q K L R N K T G K F D I T G L E E E F N K L T 630

-----+-----+-----+-----+-----+-----+-----+-----+
ACAGTCCGAATAAGGAAGGACAATATGAGAACAGGAATAAACAGTAAATATGAAAAAGCAAGAGTCTTCAGATCATTCCTATGAT
T V R I R K D N M R T G N K P V N M K K A R V L Q I I P Y D 720


```

-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTCAACAGAGTTTGTGCTGCCCAGGAAAAGGGGACAAGAATTCTCAGACTATATCAATGCTTCATACATAGATGGCTATCGACAGAAGGAT
F N R V V L P R K R G Q E F S D Y I N A S Y I D G Y R Q K D
810
-----+-----+-----+-----+-----+-----+-----+-----+-----+
IATTTTGTGCAACTCAGGGTCCTCTTGCTCATACAGTGGGAAGATTTTTGGAGAATGGTCTGGGATTGCAAGTCTCACACAATTGTAATG
Y F V A T Q G P L A H T V E D F W R M V W D C K S H T I V M
900
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTGACAGAACTGCGAGAAAAGAGAGCAGGAGAAGTGTTGCCAGTATTGGCCCTCTGAGGGATCTGTAACATATGGAACATACTCAACAGAG
L T E L R E R E Q E K C C Q Y W P S E G S V T Y G T Y S T E
990
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTGAAGAAGGACAACATTACTGATGCTTTCAGTCTCAGAGATCTCACAGTCACAAATACCACTGAAAACATAGAGAGGACAGTTCGACAA
L K K D N I T D A F S L R D L T V T N T T E N I E R T V R Q
1080
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTCCACTTCCACGACTGGCCTGAGATTGGCATTCCAGCGGAAGGTAAGGGATGATCGACCTGATTGCAGCAGTCCAGAAACAGCAGCAA
F H F H D W P E I G I P A E G K G M I D L I A A V Q K Q Q Q
1170
-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAATCCGCAATCATCCAATCATAGTGCAGTGCAGTGCTGGTGCTGGCAGAACAGGTACATTTATTGCACTGAGTAACATCCTCGAGCGA
K S G N H P I I V H C S A G A G R T G T F I A L S N I L E R
1260
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTGAAAGCTGAAGGCCTTTTAGATGTTTTCCAAACTGTTAAGAGCTTACGAATGCAACGACCGCATATGGTGCAGACATTGGAGCAATAT
V K A E G L L D V F Q T V K S L R M Q R P H M V Q T L E Q Y
1350
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GATTTCTGCTACAAGGTGGTGCAAGATTTCCCTTGACATACACTCCGATTATGCCAACTTCAAATAAATATGTTTGCCTTATA
D F C Y K V V Q D F L D I H S D Y A N F K .
1440

```

ray ryPTPR2Ab (AB033584)

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGCAACGAAGTGCAATTGTTGTGATGATGACCAAAGTGAAGAGCGATCACGGGTGAAATGTGACCAGTACTGGCCTTCCAGGGGTACA
E Q R S A I V V M M T K L E E R S R V K C D Q Y W P S R G T
90
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAAACCTTGGGATTAATCCAAGTCACTCTGCAGGATACAGTTGAACTGGCAACCTACTGTGTTAGAACATTAACATTAGACAAGATTGGC
E T L G L I Q V T L Q D T V E L A T Y C V R T L T L D K I G
180
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCCAGTGAAAAGAGGGAGGTGAGACAATTCCAATTCAGTGCATGGCCTGATCATGGGGTTCCAGAGCACCCACACCATTCTAGCCTTC
S S E K R E V R Q F Q F T A W P D H G V P E H P T P F L A F
270
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTAAGGCGAGTGAAAGCTTGCAACCTCCAGATGCTGGCCCCATGGTGGTACATTGCAGTGCCGGAGTTGGAAGGACCGGGTGTTCATT
L R R V K A C N P P D A G P M V V H C S A G V G R T G C F I
360
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTAATAGACGCCATGCTAGAGAGGATAAAGCATGAGAAAAGTGTAGATATTTATGGTTCATGTAACACTAATGAGATCACAGAGGAATTAC
V I D A M L E R I K H E K T V D I Y G H V T L M R S Q R N Y
450
-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATGGTTCAGACTGAGGACCAATATATCTTTATTCATGATGCACTGCTTGAAGCAGTGAAGTGTGGAATAACCGAAGTGCCAGCAAGAAAC
M V Q T E D Q Y I F I H D A L L E A V T C G N T E V P A R N
540
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTGTACGCGTACATCCAAAAGCTGACACAGGTGGAGCCAGGAGAAAACGTCAGTGGGATGGAAGTGGAGTTCAAGCGTCTAGCCAACACT
L Y A Y I Q K L T Q V E P G E N V T G M E L E F K R L A N T
630
-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAAGCCACACATCAAGATTCTTAGTGCCAATCTCCATGTAACAAATTTAAGAATCGCCTTGTTAATATTATGCCATATGAATCCACA
K A H T S R F I S A N L P C N K F K N R L V N I M P Y E S T
720

```



```

-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTCTGTGTCTCTACCTTCCCCTTTTGTGTCCGTCAACAGAAACACCACCTGTGCAGAGACGTCCGCTTCCCCTGAAAAGAGTTTGTGTCTCT
720
-----+-----+-----+-----+-----+-----+-----+-----+-----+
300GG
810
ray ryPTPR2Ac (AB033586)
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGCAGAGAAGTGCAACTATCGTCATGATGACAAAACCTAGAGGAAAAATCTCGGGTGAAGTGCGACCAATACTGGCCCAGCCGGGGAACA
E Q R S A T I V M M T K L E E K S R V K C D Q Y W P S R G T
180
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GATACCTATGGAATGATACAGGTGACTCTGCTGGACACCATTGAGTTGGCCACCTTCTGTGTTTCGTAATTTCTGCCTACACAAGAATGGC
D T Y G M I Q V T L L D T I E L A T F C V R N F C L H K N G
270
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCCAGTGAGAAGCGTGAGGTTTCGTTCAGTTTCAGTTCACGGCTTGGCCTGACCACGGGGTTCCCGAGTATCCTACTCCTTTCTGGCATTTC
S S E K R E V R Q F Q F T A W P D H G V P E Y P T P F L A F
360
-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTCAGAAGAGTGAAAACCTGCAATCCTCCTGACGCTGGCCCCGTGGTGGTGCAGTGCAGTGCCTGGAGTCGGTCCGACTGGCTGTTTCATT
L R R V K T C N P P D A G P V V V H C S A G V G R T G C F I
450
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTCATTGATGCCATGCTAGAGAGAATAAAACATGAGAAGACGGTTGACATCTATGGCCATGTGACACTGATGAGATCACACCGGAACACTAC
V I D A M L E R I K H E K T V D I Y G H V T L M R S Q R N Y
540
-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATGGTGCAGACAGAGGACCAGTATATTTTTATTTCATGATGCACTGCTGGAAGCAGTGGCTTGTGGGACCACTGAGGTGATGGCCAGGAAC
M V Q T E D Q Y I F I H D A L L E A V A C G T T E V M A R N
630
-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTGTA CTCTTACATCCAGAACTCACGCAGGTTGAAACAGGGGAGCATGTACAGGCATGGAAGTCAAGCGTCTGGCCAATTCC
L Y S Y I Q K L T Q V E T G E H V T G M E L E F K R L A N S
720
-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAAGCTCATACGTCCCGATTGATAAGTGCCAATCTTCCATGTAATAAGTTCAAGAACCGCCTTGTCAACATTATGCCGTATGAGACGACA
K A H T S R F I S A N L P C N K F K N R L V N I M P Y E T T
810
-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGAGTGTGTCTACAGCCAATCCGTGGGGTGAAGGATCTGACTACATCAATGCTAGCTTTATCGATGGATACAGGCAACAGAAAGCATAC
R V C L Q P I R G V E G S D Y I N A S F I D G Y R Q Q K A Y
900
-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATTGCCACACAGGGTCTTTAGCAGAGACTACAGAGGACTTCTGGAGAATGTTATGGGAGAACAACTCCACCATTGTTGTCATGCTGACC
I A T Q G P L A E T T E D F W R M L W E N N S T I V V M L T
990
-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAACTGCGAGAAATGGGCCGTGAGAAGTGTACCAGTACTGGCCAGCTGAACGATCCGCCAGGTATCAGTACTTTGTTCGTGGATCCGATG
K L R E M G R E K C H Q Y W P A E R S A R Y Q Y F V V D P M
1080
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCAGAGTACAACATGCCACAATACATCCTCCGTGAGTTCAAAGTCACAGATGCAAGGGATGGCCAGTCACGGACTGTCCGGCAGTTCCAG
A E Y N M P Q Y I L R E F K V T D A R D G Q S R T V R Q F Q
1170
-----+-----+-----+-----+-----+-----+-----+-----+-----+
PTTACAGACTGGCCAGAGCAAGGAGTGCCAAAATCCGGCGAGGGGTTTCATTGACTTTATTGGGCAAGTGCATAAACTAAGGAACAGTTT
F T D W P E Q G V P K S G E G F I D F I G Q V H K T K E Q F
1260
-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGCCAAGATGGTCCCATCTCAGTTCACTGCAGTGCCTGGAGTTGGCCGACTGGGGTCTTCATCACGCTCAGCATTTGCTTGGAAAGAATG
G Q D G P I S V H C S A G V G R T G V F I T L S I V L E R M

```

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGGTACGAAGGGGTGGTGGACATTTTCCAGACTGTCAAGATGCTCCGAACACACGCGACCAGCCATGGTACAGACAGAGGATGAATATCAG
R Y E G V V D I F Q T V K M L R T Q R P A M V Q T E D E Y Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTCTGCTATCAAGCAGCATTGGAGTACCTTGGGAGTTTTGATCATTATGCAACGTAAAAAGCCATTGATCATCAAACCTAACAGGCCACA
E C Y Q A A L E Y L G S F D H Y A T .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATGGACACGGAGCACCTCTTCTGAGCTACATAGTGTGCTTGAAAAGTACTTTAATTCTAACCACCGCCGGGGGAAAGA

```

ray ryPTPR2B (AB033587)

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAAGAACAATCAGTTTGTATCGTTATGGTCACAAATCTGGTGAAGTTGGAAGAGTCAAATGTTACAAATATTGGCCAGATGATAGTGAA
Q E Q S V C I V M V T N L V E V G R V K C Y K Y W P D D S E
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTTTATGCAGATTTCAAAGTTACATTCGTGGAAGAGGAGCCTCTTGCAGAAATATGTTGTTCGAACATTTACATTGGAAAGGAGAGGCTTT
V Y A D F K V T F V E E E P L A E Y V V R T F T L E R R G F
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GATGAAATTCGTGAAGTGAAGCAGTTCATTTTCATTGGCTGGCCTGATCATGGTGTTCCTTATCATGCAACTGGTCTACTAGCATTATC
D E I R E V K Q F H F I G W P D H G V P Y H A T G L L A F I
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGCCGTGTGAAATTCTCAAATCCCCCAGGAGCAGGGCCAGTAGTTGTACATTGCAGTGCTGGGGCTGGAAGGACTGGCTGTTATATTGTT
R R V K F S N P P G A G P V V V H C S A G A G R T G C Y I V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATAGATATCATGTTAGACATGGCAGAGCGAGAAGGAGTTGTTGATATCTATAATTGTGTTAAAGCATTACGATCAAGACGCATTAATATG
I D I M L D M A E R E G V V D I Y N C V K A L R S R R I N M
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTCAGACAGAGGAACAATATATATTCATTTCATGATGCCATTCTGGAAGCTTGTCTGTGTGGGGACACATCAATACCAGTCTGCGAAATC
V Q T E E Q Y I F I H D A I L E A C L C G D T S I P V C E I
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAGCAAGCCTATTACGAAATGATACGAGTTGATTACAGAGTAACCTCTTCCAGCTTAAAGATGAATTTCAAACCCTAAATTCTGTATCA
K Q A Y Y E M I R V D S Q S N S S Q L K D E F Q T L N S V S
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCACAACTACAACCCGAAGACTGCAGCATTGCTCTTCTACCCAGGAATCATGAGAAGAATCGATACATGGACGCATTGCCACCAGACAGA
Q L Q P E D C S I A L L P R N H E K N R Y M D A L P P D R
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGCCTGCCATTTTTTAATTACCATAGATGGTGAAAGCAGTAACCTACATTAATGCTGCTCTAATGGATAGCTATAGGCAACCAGCTGCTTTT
C L P F L I T I D G E S S N Y I N A A L M D S Y R Q P A A F
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATAGTGACACAGCATCCATTGCCAAACACAGTCAAAGATTTCTGGAGACTGGTCTATGATTATGGCTGCACTTCTGTTGTTCATGCTGAAT
I V T Q H P L P N T V K D F W R L V Y D Y G C T S V V M L N
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGTTGGACCTTGTTTCAGGGATGCCCGCAGTATTGGCCAGAAGAGGGGACACTTCGATATGGGCCTATTCAAGTTGACTGCGTTTCGTGC
E L D L V Q G C P Q Y W P E E G T L R Y G P I Q V D C V S C
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCAATGGATTGTGATCTAGTAAGTCGGCTCTTCAGGATCTGCAATCTTACAAGGCCACAGGAAGGTTATCTCATGGTGCAGCAGTTCCAG
S M D C D L V S R L F R I C N L T R P Q E G Y L M V Q Q F Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TACTTGGGATGGGCAGCTCACAGGGAAATACCTGCCTCAAAACGATCATTTCTGAAGTTAATTCTCCAGGTTGATAAATGGCAAGAAGAA
Y L G W A A H R E I P A S K R S F L K L I L Q V D K W Q E E

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GAGCACAATGTGAGTGTCTATTGTGATGATCACCAACCTGGTAGAAAAAGGAAGGAGGAAATGTGATCAGTACTGGCCATCAGAAAAAGT
 E H N V S V I V M I T N L V E K G R R K C D Q Y W P S E N S 90
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GAAGAATATGGACACATCCTTGTATCGCTGAAGAACACAAAAGTTTCATCTTATTACACTGTTAGACAGTTTACTGTACGCAACACAAAG
 E E Y G H I L V S L K N T K V S S Y Y T V R Q F T V R N T K 180
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TTTAAAAAGGGATCATTGAAAGGAAGGCAGAATGATCATATAGTGACCCAGTATCATTACACCCAGTGGCCTGATATGGGTGTTCCAGAG
 F K K G S L K G R Q N D H I V T Q Y H Y T Q W P D M G V P E 270
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TATGCCCTTCCTGTTTTAACATTTATTCGACACTCAGCAGCTGCTAAAGCCAGTCATACTGGGCCTGTTGTGGTTCATTGCAGTGCTGGT
 Y A L P V L T F I R H S A A A K A S H T G P V V V H C S A G 360
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GTTGGAAGAACTGGCACATACATTGTAATAGACAGCATGTTGCAACAAATCAAACAGAAAGGCACTGTCAATGTGCTGGGATTCCCTTAAA
 V G R T G T Y I V I D S M L Q Q I K Q K G T V N V L G F L K 450
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CACATCAGAACACAAAGGAACTATTTAGTGCAAACAGAGGAGCAGTATATCTTCATTCATGATGCCCTTGTGGAAGCCATCCTTACTAAG
 H I R T Q R N Y L V Q T E E Q Y I F I H D A L V E A I L T K 540
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GACACTGAGGTTTCCTGCCAATCATCTGCATACCTACGTAAATGCTCTTCTTACTCCGGGACCATCGGGAAAAACAAAAGTCTAGAGAAACAA
 D T E V P A N H L H T Y V N A L L T P G P S G K T K L E K Q 630
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TTCAAGCTGGTGAGTCAGCCTCAGGCCAAGCAGTGTGATTATTCTGCAGCCCTGAAGCAATGCAATCGAGACAAGAACAGAAATTCCTCC
 E K L V S Q P H A K Q C D Y S A A L K O C N R D K N R N S S 720


```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATCATTCCTGTGGAGAGATCAAGAGTTGGACTTGCAACACTGACTGCTGAGGGCTCTGATTACATCAATGCATCATATGTTTCGGGATT
I I P V E R S R V G L A T L T A E G S D Y I N A S Y V S G F
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CATCGAATTGATGAGTTCATCGTCAACCAACATCCACTGCCATACACTACAAAGGACTTCTGGCGGATGATATGGGACTACAATGCTCAA
H R I D E F I V T Q H P L P Y T T K D F W R M I W D Y N A Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATTATAGTCATGCTGCCAGATAACCAAGGACTGACCCCATATTGGCCAAAGAAGGATGGCCCAATGAACATGAAGCATTTACTGTGAGT
I I V M L P D N Q G L T P Y W P K K D G P M N Y E A F T V S
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTACTCAATGAGAGCCACCTGTGTCTTTCCAATGAAGAGCAGCTTACCATCAAGGATTTTCTTCTGGAAGCAACACAGGATGACTATGTA
L L N E S H L C L S N E E Q L T I K D F I L E A T Q D D Y V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTAGAAGTAAGACACTACCAGTCTCCAAGGTGGCCAAATCCAGATGGTCCAATAAGCAAAACCTTTGAGCTCCTCAATTTGATACGGGAT
L E V R H Y Q S P R W P N P D G P I S K T F E L L N L I R D
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCTGCTGCTACTAGAGATGGGCAATAATCGTTTACGATGAGCATGGAGGTATTACAGCTGGCACCTTCTGTGCTTTGACCAAATTGATG
A A A T R D G P I I V H D E H G G I T A G T F C A L T K L M
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GATCAGCTGGAGAGTGAAAACCTCAGTGGATGTTTACCAAGTGGCTAAAATGATCAATTTAATGAGGCCTGGAATATTCACCGATATTGAC
D Q L E S E N S V D V Y Q V A K M I N L M R P G I F T D I D
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGTTTCAATTTTTATACAAAGCTATACTCAGCCTGGTCAACTCAAGGGAAGACAAAAACACTGCAGCTATACTGGAACCAATGGTCTT
Q F Q F L Y K A I L S L V N S R E D K N T A A I L E T N G L
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GCTCTGTCTGCTGAAGATTACAATGGGGCAGAAAGTATGGAGTCTCTGGTGTAAATAGTGAAGTGGATGTTTTTACCAACTCAGAATCCTG
A L S A E D Y N G A E S M E S L V .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTTGATCTTGTTAGCTATACAAGTATTCACCTTGAATGCTAATATCCATTTTATGTCATCATTGGATGGTGACCCGTGGATGACATAAT
ATAAAAAAAAAA

```

ray ryPTPR4b (AB033589)

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGCAAATACTGCAACAATTGTCATGGTCACAAATCTGAAAGAGCGGAAAGAGTGCAAATGTGCGCAGTACTGGCCAGACCAGGGTTGT
E Q N T A T I V M V T N L K E R K E C K C A Q Y W P D Q G C
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGGACATACGGGAACATTCGAGTCTTCTGTAGAAGACATGATGGTGTAGTGGATTATACCATTGCAAGTTCTGTATTACAGCAAGTAGGA
W T Y G N I R V S V E D M M V L V D Y T I R K F C I Q Q V G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GATGTGATGAACAAGAAACCTCAGAGGCTGGTTACTCAGTTCCACTTTACAGTTGGCCTGATTTTGGAGTACCATTCACTCCAATTGGC
D V M N K K P Q R L V T Q F H F T S W P D F G V P F T P I G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATGCTGAAATTTCTGAAGAAAGTGAAGAACTTGTAAATCCATCCTTTGCAGGTGCAATCGTTGTGCACTGCAGTGTGTTGGTTCGTACA
M L K F L K K V K T C N P S F A G A I V V H C S A G V G R T
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGCACTTTCATTGTTATTGATGCCATGTTGGATATGATGCATGCAGAGAGGAAGGTGGACGTATTTGGATTTGTGTCACGGATTTCAGCT
G T F I V I D A M L D M M H A E R K V D V F G F V S R I R A

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGAGGTGTCAAATGGTTCAGACAGATATGCAGTATGTATTTATATATCAAGCACTATTGGAACATTATTTATATGGAGACACAGAAGTGC
Q R C Q M V Q T D M Q Y V F I Y Q A L L E H Y L Y G D T E L
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 540
GAAATGCACTTCACTTGAAAGCCATTTGCAAAAGCTTTACAATAGAGTATCTGGATCAGGCTGCAATGGTCTAGAGGAAGAATTCAAGAAA
E V T S L E S H L Q K L Y N R V S G S G C N G L E E E F K K
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 630
TTGACATCGATCAAAATTCAGAACGATAAGATGAGAACGGGTAAATTTGCCAGCCAACATGAAGAAAAACAGAGTTCTTCAGATCATCCCA
I T S I K I Q N D K M R T G N L P A N M K K N R V L Q I I P
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 720
TATGAATTTAACCAGTGCATCATTCCTGTAAAGAGAGGTGAAGAAAACACAGATTATGTAAATGCCTCTTTTCATTGATGGATATCGACAG
Y E F N R V I I P V K R G E E N T D Y V N A S F I D G Y R Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 810
AAAGATTCATACATAGCCAGTCAAGGACCTCTTCAACATACTATAGAAGATTTCTGGAGAATGATTTGGGAATGGAAGTCGTGTTCCATT
K D S Y I A S Q G P L Q H T I E D F W R M I W E W K S C S I
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 900
GTTATGTTAACAGAATTAGAAGAAAGAGGCCAGGAGAAATGTGCACAGTATTGGCCATCAGAAGGAAGTGTAACTTACGGTGATATTACA
V M L T E L E E R G Q E K C A Q Y W P S E G T V T Y G D I T
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 990
ATTGAAATAAAAAAGGAAGAAGAAAGTGAAAGTTACACAGTGCCTGATTACTGGTGACTAACACACGGGAAAACAAAGTTCGTGAGATC
I E I K K E E E S E S Y T V R D L L V T N T R E N K V R Q I
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1080
AGACAGTTTTCATTTTCATGGTTGGCCTGAGGTGGGTATTCCTGCTGACGGGAAGGGAATGATCAACCTGATTGCAGCAGTGCAGAAACAA
R Q F H F H G W P E V G I P A D G K G M I N L I A A V Q K Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1170
CAGCAGCAGTCAGGAAACCATCCAATCACCGTGCAGTGCCTGGAGCCGGAAGAACAGGAACCTTCTGTGCCCTGAGTACTGTTTTG
Q Q Q S G N H P I T V H C S A G A G R T G T F C A L S T V L
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1260
GAAAGGGTGAAGGCTGAGGGTATCTTGACGTATTTAGACAGTAAAGAGTCTACGGCTACAGAGACCACATATGGTCCAGACACTGGAA
E R V K A E G I L D V F Q T V K S L R L Q R P H M V Q T L E
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1350
CAGTATGAATTTCTGCTACAAGGTTGTGCAAGAATACATTGATGCATTTTCGGATTATGCTAACTTCAAATGAAAAATGGCCTTCGTAATA
Q Y E F C Y K V V Q E Y I D A F S D Y A N F K .
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1440
TTTTGTAATATTCTATTTGTTAATATACTGAAACCCGTGTACATATCTTATTGGTTTTAGAGGTTGGTACCATAAGGCTTTCATATTAGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1530
TTGAGATTTTTTATGTGCGAGATGGTTAGCATATTAGTCTCAGTTAAGCAATTTTATTGAGTAATGTGGCCTGATGTTTGGATTAATGTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1620
GATGAGATTTATTTAGTAGCTGCTTCAAAGGGTATTCATTTGAATTCAAAGGAGGAAGGACAGAATTTAAAAAAA
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1710

ray ryPTPR5b (AB033590)

-----+-----+-----+-----+-----+-----+-----+-----+-----+ 90
GAACAAAATACAGGAATCATCGTTATGATTACAAATTTAGTGGAAAAAGGAAGACGAAAATGTGATCAGTACTGGCCCAGTGAGAACAAT
E Q N T G I I V M I T N L V E K G R R K C D Q Y W P S E N N
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 180
GAAGAGTATGGCAACATAATTGTCCACCTGAAGAGTACAAAAGTAAGGGCATGCTGTACTGTACGTCACTTCACCATACGGAACACAAAA
E E Y G N I I V T L K S T K V R A C C T V R H F T I R N T K

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTAAAAAGGGGCAAAAGGGGAAACCCCAAGGGACGTGAGAATGAACGCATGGTTATCCAGTATCACTACACTCAGTGGCCCGACATGGGG
V K K G Q K G N P K G R Q N E R M V I Q Y H Y T Q W P D M G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTACCAGAGTATGCACTTCCTGTACTCACATTTGTAAGAAAATCATCTGCAGCTCAAACACCAGACATGGGACCAGTGGTGGTACACTGC
V P E Y A L P V L T F V R K S S A A Q T P D M G P V V V H C
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGTGCTGGTGTGGTAGAACTGGAACATACATTGTGCTTGACAGTATGTTGAAACAGATAAAGGATAAAAGCACAGTTAATGTTCTAGGT
S A G V G R T G T Y I V L D S M L K Q I K D K S T V N V L G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTCTCAAACATATCAGGACCCAGAGGAACTATCTTGTCCAAACAGAAGAACAATACATTTTATTTCATGAAGCTTTGCGGGAAGCAATA
F L K H I R T Q R N Y L V Q T E E Q Y I F I H E A L R E A I
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CATGGAAAAGAAACAGAAGTACATTCCTCTCAGCTGCATGGTTACGTAAACAGTATTTTAAATTCAGGTCTTTCTGGGAAGATGAGGCTG
H G K E T E V H S S Q L H G Y V N S I L I P G L S G K M R L
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGAAGCAGTTCAAGTTGGTAACCTCAGTGTAATGCAAAATTTATAGAATGCTTCAGTGCCAGAAAGAATGCAACAAAGAGAAGAACAGG
E K Q F K L V T Q C N A K F I E C F S A Q K E C N K E K N R
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AACTCCTCAGTTGTGCCAACTGAACGAGCTCGGGTAGGCCTCGCACCATTACCTGGGACCAAAGGAACTGATTATATCAATGCCTCCTAT
N S S V V P T E R A R V G L A P L P G T K G T D Y I N A S Y
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATCATGGGGTACTACAGAAGCAATGAATTTATTATTACTCAACATCCTTTGCCACACACAACGAAGGACTTTTGGCGAATGATTGGGAC
I M G Y Y R S N E F I I T Q H P L P H T T K D F W R M I W D
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CACAATGCACAAATTATTGTCTGTTACCAGATAACCATGGACTTGCTGAGGATGAATCTGTGTACTGGCCAAGCCGAGAAGAATCCATG
H N A Q I I V M L P D N H G L A E D E S V Y W P S R E E S M
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AATTGTGAAGCTTTTACTGTTACCCTCATCAGCAAAGACATATTGTGCCTCTCTAATGAAGAACAAGTTATCATCCATGATTTTCATCCTG
N C E A F T V T L I S K D I L C L S N E E Q V I I H D F I L
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAGGCAACACAGGATGATTATGTACTGGAAGTGGCTCACTTTCAGTGTCCAAATGGCCCAATCCAGATGCTCCACTGAGTAGCACCTTT
E A T Q D D Y V L E V R H F Q C P K W P N P D A P L S S T F
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AACTTATTAACGTATTAAGAAGAGGCCTCAACCAGAGATGGACCAACCATTGTTACAGATGAGTTTGGAGGCGTCTCTGCTGGGCTG
L I N V L K E E A S T R D G P T I V H D E F G G V S A G L
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTCTGTGGCCTTACCACCTTTACAGCAGTTGGAGTGTGAATCGGCTGTGGATGTGTACCAGGTTGCAAAGATGATAAACCTTATGAGG
F C G L T T L S Q Q L E C E S A V D V Y Q V A K M I N L M R
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCAGGAGTTTTACAGACATTGAGCAGTATCAGTTCCTATATAAAGCAATGCTCAGCTTAATCAGCACAAAAGAGAATGGGAATGGCCCC
P G V F T D I E Q Y Q F L Y K A M L S L I S T K E N G N G P
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCATCATTTGACAAAAACGGCACCATCATTATCACTGATGAATCAGATCCAGCAGAGAGCATGGAATCTCTTGTGTAATATGAAGCTTTT
S S F D K N G T I I I T D E S D P A E S M E S L V .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGCAAATGGCACTTCATTTGAAAATGTTTAAGGACTGAATGAACCTTTTGAGGCCTTTTATTGCCAGGTTCCAATTGTACTGACAACCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACTGTACTTCACACAG

```

90
+-----+-----+-----+-----+-----+-----+
GACACTACAGCAGCAGGAGTGC AAGCTTC TTTATAGTCGAAAGGATGGGCAACGGCAAGAAAATAAATATAAGAACC GGTTACAAGAACAT
T L Q Q Q E C K L L Y S R K D G Q R Q E N K Y K N R Y K N I
180
+-----+-----+-----+-----+-----+-----+
CCTGCCTTTTGACCATACTCGAGTAGTCCTACATGATGGAGATGTGACTGAACCTGGTTCTGATTACATCAATGCTAATTTTCATTTTGCA
L P F D H T R V V L H D G D V T E P G S D Y I N A N F I L H
270
+-----+-----+-----+-----+-----+-----+
TGATATTGACAGCAATCGAACCACAACCAAGCCAAAAAAGACTTACATTGCCACACAGGGATGTTTGCATGACACAGTGAATGACTTCTG
D I D S N R T T T K P K K T Y I A T Q G C L H D T V N D F W
360
+-----+-----+-----+-----+-----+-----+
GAGGATGGTTTTTCAAGAAA ACTCACGAGTAATTGTCATGACAACGAAAGAGTGGAAAGAGGAAAGAGCAAATGTGTGAAATACTGGCC
R M V F Q E N S R V I V M T T K E V E R G K S K C V K Y W P
450
+-----+-----+-----+-----+-----+-----+
AGATGAAATGTCCTCAAAGGAATATGGTGCAATGCGTGTCAAGAATGTTCTGTGAGAACCCAGCACATGACTATATTTTAAGGGAGCTCAA
D E M S S K E Y G A M R V K N V R E N P A H D Y I L R E L K
540
+-----+-----+-----+-----+-----+-----+
ACTTTCTAAAGTAGGACAGGGGAGTACTGAAAGGACGGTGTGGCAATATCATTTTAAAACATGGCCAGATCACGGTGTCCCAGTGATCC
L S K V G Q G S T E R T V W Q Y H F K T W P D H G V P S D P
630
+-----+-----+-----+-----+-----+-----+
AGGTGGTGTCTCTGGATTTCCCTTGAAGAAGTTAATATCAAACAGGAGAATATCCCAGAAGCTGGGCCAATCGTGGTCCACTGCAGTGCAGG
G G V L D F L E E V N I K Q E N I P E A G P I V V H C S A G
720
+-----+-----+-----+-----+-----+-----+
GATTGGACGCACAGGAACGTTCA TTGTGATTGACATCCTGATTGACATCATCAGAGAGAAAGGTGTGGACTGTGACATCGATGTTCCAAA
I G R T G T F I V I D I L I D I I R E K G V D C D I D V P K
810
+-----+-----+-----+-----+-----+-----+
GAGTATCCAGATGGTGCCTCACAACGTT CAGGGATGGTTTCAGACAGAAGCCAGTACAGATT CATTTACATGGCTGTT CAGCACTACAT
S I Q M V R S Q R S G M V Q T E A Q Y R F I Y M A V Q H Y I
900
+-----+-----+-----+-----+-----+-----+
CGAAACACTGCAGCGTCGGATCGAAGAGGAGCAGAAAAGCAAGATAAAAGGCCGTGAATACACAAACATAAAATATTCCTTATCGGACCT
E T L Q R R I E E E Q K S K I K G R E Y T N I K Y S L S D L
990
+-----+-----+-----+-----+-----+-----+
GACGTGTACAGGAGATCAGAGTCCTTTCCCTCCTTGCACACCCAATCAAGTGTGTACAGAGATGAAAGATGAGGGTGGCAGGATATATGA
T C T G D Q S P F P P C T P N Q V C T E M K D E G G R I Y E
1080
+-----+-----+-----+-----+-----+-----+
AAATGTTGGCTTGATGCGACAGCAA AAGGGCCATAGATGAAAGTGTCCAGTGTGTGTAACAGAATGGAAAG
N V G L M Q Q Q K G H R .

+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GCTTCTACAGCAACAAGAATATAAACTGTTGTACAGCAGGAAGGAAGGCCAGAGAATAGAAAATAAAAGCAAAAACAGATACAAAATAT
 L L Q Q Q E Y K L L Y S R K E G Q R I E N K S K N R Y K N I
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 90
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TTACCATTTGATAATACCAGAGTTAGGCTGAAGGAAGTTGATGAAAGTGTAGTGGGATCTGATTATATCAATGCTAATAACATCACAAA
 L P F D N T R V R L K E V D E S V V G S D Y I N A N N I T N
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 180
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 TGGAAAAATGGAGATGATTGTAAAACTACATTGCAACACAAGGCTGTCTTCAGAACACCATCAATGATTTTTGGAAAAATGATTACCA
 G K N G D D C K N Y I A T Q G C L Q N T I N D F W K M I Y Q
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 270
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GAAAAATGCTCATGTTATTGTTCATGACAATAAGAAATAGAAAGAGGAAAGTGTAAAGTGCCTTGCCTACTGGCCAGATGTCAAACCTTC
 E N A H V I V M T T K E I E R G K C K C V R Y W P D V K T S
 +-----+-----+-----+-----+-----+-----+-----+-----+-----+ 360

```

-----+-----+-----+-----+-----+-----+-----+-----+
AACAGAGTACGGCATGATATCAGTGAGGAATCTGGAAGAACGCATGGCACAAAGATTATATTGTACGAGACCTGGAGGTTATGCGACTAGA
T E Y G M I S V R N L E E R M A Q D Y I V R D L E V M R L D
450
-----+-----+-----+-----+-----+-----+-----+-----+
CTGGAATGAACGTCCTCGGCACGTGTGGCACTACCAGTACCTCAGCTGGCCTGACCACGGGGTTCCCAATGAACCAGGAGGAGTTTAAAG
W N E R P R H V W H Y Q Y L S W P D H G V P N E P G G V L S
540
-----+-----+-----+-----+-----+-----+-----+-----+
CTTCTTGGACCAAGTAAACAGAACACAGCAAAGTATCCCTGACACTGGGCCCATATAGTGCAGTGCAGTGCCTGGAATAGGGAGAACAGG
F L D Q V N R T Q Q S I P D T G P I I V H C S A G I G R T G
630
-----+-----+-----+-----+-----+-----+-----+-----+
AACTATAATCGTCATAGATATGTTAATATCTGACATCAGCAGACAAGGTTTGGACTGTGATATAGACATCCCAAAAACGATTCAAATAGT
T I I V I D M L I S D I S R Q G L D C D I D I P K T I Q I V
720
-----+-----+-----+-----+-----+-----+-----+-----+
TCGAAAACAACGATCAGGAATGGTACAAACTGAAGCTCAGTACAAGTTCATATACATGGCTGTACAGCAGTATATTGAAACAGTGCAGAG
R K Q R S G M V Q T E A Q Y K F I Y M A V Q Q Y I E T V Q R
810
-----+-----+-----+-----+-----+-----+-----+-----+
GAGACTACAGGAAGAACAGAGAAGCAAAATAAAGGACCGAGAGTATTCAAATATCCGATATCCTCCTATGGAGGTTAAGACCAAAAGACA
R L Q E E Q R S K I K D R E Y S N I R Y P P M E V K T K R H
900
-----+-----+-----+-----+-----+-----+-----+-----+
TGCTGGCACGTCTCGAGATCTTTTCATTGTTGAATGAAGAATCACATTGTCTTTATGAAAATTTAAACATCAAAACTTCCAAAGCTTTAGG
A G T S R D L S L L N E E S H C L Y E N L N I K T S K A L G
990
-----+-----+-----+-----+-----+-----+-----+-----+
AGGCAGCAACACAAAACGATAAAAGCGTCAAGTGGACTGGCTCCGTCACCCCTACCGGAACACAGACTGCTGACAGCTGCTCACCAGAGT
G S N T K R .
1080
-----+-----+-----+-----+-----+-----+-----+-----+
CCTCAGTGCTGAGCCAGTCTTTCAAGTTGCCCCAGGTGTAGACCATCTTACATCTTGCAGGTGAAGATCCATACACTCCCAGGGATGAG
1170
-----+-----+-----+-----+-----+-----+-----+-----+
GCTTTTGGAGTTTCTGTTGGTGTTTTTGTAGCTCTGAGTTATTACGGGATGAGATTGCTAGTCATATGCCAAACTATCCTCCTTTTCAG
1260
-----+-----+-----+-----+-----+-----+-----+-----+
CCAGGCT
1350

```


sponge sPP2B

sponge sPP1a

-----+-----+-----+-----+-----+-----+-----+-----+90
GAAGCAATCAATGGAGACCATATGCCTGCTTTTGGCTTACAAGATACGATATCCGAAAACCTTCTTTCTGCTCAGAGGAAACCACGAGTG
K Q S M E T I C L L L A Y K I R Y P E N F F L L R G N H E C

```

-----+-----+-----+-----+-----+-----+-----+-----+
CGCGAGCATCAACAGGATTTACGGGTTTTACGACGAGTGCAAGCGTCGCTACAACATACGTCTATGGAAGACGTTACAGACTGCTTCAA
  A S I N R I Y G F Y D E C K R R Y N I R L W K T F T D C F N
-----+-----+-----+-----+-----+-----+-----+-----+
270
TTGTCTACCAGTAGCGGCCGTATTAGACGAAAAGATTTTCTGCTGTCTGAGGAGTCTCGCCTGACCTTCAGTCCACGGAGCAAATTCG
  C L P V A A V L D E K I F C C H G G L S P D L Q S T E Q I R
-----+-----+-----+-----+-----+-----+-----+-----+
360
ACGCATAATGCGTCCGACAGATGTCCCCGACCAGGGCCTTCTGTGCGATCTGTTGTGGGCAGACCCCGATAAGGAGAACGCAGGATGGGC
  R I M R P T D V P D Q G L L C D L L W A D P D K E N A G W A
-----+-----+-----+-----+-----+-----+-----+-----+
450
CGAAACGACAGGGGCGTCTCCTTTACATTTCGGAACGGACATCATAAATAAATTTCTGAACAAACACGAATTCGATCTTATATGCCGCGC
  E N D R G V S F T F G T D I I N K F L N K H E F D L I C R A
-----+-----+-----+-----+-----+-----+-----+-----+
540
GCATCAGGTGGTTGAAGATGGGTACGAGTTTTTTTCAAACGTCGAATGGTCACTTTATTTTCCGCGCCCACTACTGTGGAGAATTCTGA
  H Q V V E D G Y E F F S K R Q L V T L F S A P N Y C G E F D
-----+-----+-----+-----+-----+-----+-----+-----+
630
CAACGCAGGAGCAATGATGAGTGTGATGAGTCGCTCATGTGCTCCTTCCAGATAATCAAGCCACCGGAAAAAGAAAGCAAAAGTTCTGGGG
  N A G A M M S V D E S L M C S F Q I I K P P E K K A K F W G
-----+-----+-----+-----+-----+-----+-----+-----+
720
TAGCGCTAAAAGTACAAGCATAGATGGAGCAAAACCAGGTGACTCTAAGAAGAAGAAGTGATGCAGTTATAGGCATTGTGTGAATTCGT
  S A K S T S I D G A K P G D S K K K K .
-----+-----+-----+-----+-----+-----+-----+-----+
810
CAGACTCGTCTTATGAATGTTCTTCTGTTCCATTTCAATTTCCACACTTAAGGTTATTTTAAATTCTTAATTCATTTCTTTTATTTT
-----+-----+-----+-----+-----+-----+-----+-----+
900
ATTTTTTTAATACCA

```

sponge sPPlc

```

-----+-----+-----+-----+-----+-----+-----+-----+
90
GAAGCAGTCGTTGGAACCATCTGTCTACTGCTGGCCTACAAGATCAAGTACCCAGAGAAGTTCTTTCTTCTCCGAGGCAACCATGAGTG
  K Q S L E T I C L L L A Y K I K Y P E N F F L L R G N H E C
-----+-----+-----+-----+-----+-----+-----+-----+
180
TGCAAGCATCAATAGAATATATGGCTTCTATGATGAGTGTAAGAGACGGTACAGTATTAAGATCTGGAAGACATTCACAGATTGTTTCAA
  A S I N R I Y G F Y D E C K R R Y S I K I W K T F T D C F N
-----+-----+-----+-----+-----+-----+-----+-----+
270
CTGCCTGCCGATTGCTGCCGTCAATTGACAACAAGATATTCTGCTGCCATGGAGGATTATCACCTGATCTTCATGACTTCGAACAAATCAG
  C L P I A A V I D N K I F C C H G G L S P D L H D F E Q I R
-----+-----+-----+-----+-----+-----+-----+-----+
360
ATCGTTGCCAGGCCTACGGATGTCCAGACACAGGACTTCTGTGACCTCCTCTGGTCCGACCCAGAAAAGGAGATTCAAGGCTGGGG
  S L P R P T D V P D T G L L C D L L W S D P E K E I Q G W G
-----+-----+-----+-----+-----+-----+-----+-----+
450
CGAGAACGATCGTGGAGTGTCTTCACTTTCGGAATGACATTGTCACCAAGTTTCTCAATAGACATGATTTGGATCTTATTTGCAGAGC
  E N D R G V S F T F G N D I V T K F L N R H D L D L I C R A
-----+-----+-----+-----+-----+-----+-----+-----+
540
ACATCAGGTTGTGCGAGGATGGTTACGAATCTTTGCTCGTCGTCAGTTAGTAACCTTGTTTTCTGCCCCTAACTACTGCGGTGATTATGA
  H Q V V E D G Y E F F A R R Q L V T L F S A P N Y C G D Y D
-----+-----+-----+-----+-----+-----+-----+-----+
630
CAATGCTGGAGGAATGATGACAGTGGATGAAAACCTCTTGTTTCATTCCAGATTTTGAAACCTTCAGAGAAGAAAACCTAAGTACCAGTA
  N A G G M M T V D E N L L C S F Q I L K P S E K K T K Y Q Y

```

```

-----+-----+-----+-----+-----+-----+-----+
TTCAGGATTGAACTCTGGGAAGCCAACCACAACATCATCTAAGGCTGCCACGATAAAGGCTCACTGATCTTCATCTCCTGCCTTCTGCCT
S G L N S G K P T T T S S K A A T I K A H .
-----+-----+-----+-----+-----+-----+-----+
TCTGATGTATCTACTCTTACCATACTGTTAGCTGTGTGAATATACCAGTACTGACTTATGATAGAGGCATTTCATTGTTATTGTGTGTTGT
-----+-----+-----+-----+-----+-----+-----+
GTTTTTTGGTGTACGCATCTCTGCGGCTTATTGTGTCGTTTGATTTATTCTGT

```

sponge sPP2A

```

-----+-----+-----+-----+-----+-----+-----+
CTACTACTCTGTGGAACAGTATCACTTCTTGTGCACTAAAAGTGCGGTTTCCAAACAGAGTTACGATCCTCAGAGGAAATCACGAGAG
Y Y S V E T V S L L V A L K V R F P N R V T I L R G N H E S
-----+-----+-----+-----+-----+-----+-----+
CAGGCAGATTACGCAAGTGTATGGGTTTTACGATGAGTGTAAAGGAAATATGGAATCCAAATGTGTGGAAGTATTTCACTGACCTTTT
R Q I T Q V Y G F Y D E C L R K Y G N P N V W K Y F T D L F
-----+-----+-----+-----+-----+-----+-----+
CGACTATTTACCACTAACGGCATTGGTGGATGGTCAGATATTCTGTCTGCATGGTGGTCTATCACCTTCTATTGACACTCTTGATCATAT
D Y L P L T A L V D G Q I F C L H G G L S P S I D T L D H I
-----+-----+-----+-----+-----+-----+-----+
CAGATCATTGGATCGATTGCAAGAAGTCCCTCATGAGGGTCCAATGTGTGACCTGTTGTGGTCAGATCCTGATGATCGTGGAGGCTGGGG
R S L D R L Q E V P H E G P M C D L L W S D P D D R G G W G
-----+-----+-----+-----+-----+-----+-----+
CATCTCACCTCGTGGGGCTGGCTACACATTTGGGCAAGACATCTCGGAAAACCTCAACCACACAAATGGTTTGACACTGATCAGCAGGGC
I S P R G A G Y T F G Q D I S E N F N H T N G L T L I S R A
-----+-----+-----+-----+-----+-----+-----+
TCACCAGCTGGTGATGGAGGGTTATACTGGTGTGATGAACGCAATGTAGTCACAATTTTCAGTGCACCCAACTATTGCTACCGTTGTGG
H Q L V M E G Y N W C H E R N V V T I F S A P N Y C Y R C G
-----+-----+-----+-----+-----+-----+-----+
AAACCAGGCAGCCATCATGGAGTTGGATGATGGGCTGAAATACACATTTTTGCAGTTTGATCCAGCTCCCAGAAGAGGGGAACCACTTGT
N Q A A I M E L D D G L K Y T F L Q F D P A P R R G E P L V
-----+-----+-----+-----+-----+-----+-----+
AACACGACGCATTCCGGATTACTTCATGTAGATTCTTCTCTCACATTCCTTTTTTTGGGACATTGTCCAACCTAAGTGATACCCTGAACA
T R R I P D Y F M .
-----+-----+-----+-----+-----+-----+-----+
CATTGTTAGTGCAGCTGTTGTGGTTAAATGTTAACCCCATATTTTGCACCTATATATAGTTGTAGGAGGCATAGAATTATTTTGAAGCT
-----+-----+-----+-----+-----+-----+-----+
GTTTTGATGTGTTGCTGGTTAGTGTGTTCAATTTATATTATACATGTACAGTGAATATTCAATTGGTAACGAT

```

sponge sPP5

```

-----+-----+-----+-----+-----+-----+-----+
AGAGGGTCATTCTCTGTTGAGTGTATCCTGACTCTTTTCGGCTTCAAGTTATTGTACCCTAACCACTTCTTTTGGCGCGAGGTAACCAT
R G S F S V E C I L T L F G F K L L Y P N H F F L A R G N H
-----+-----+-----+-----+-----+-----+-----+
JAAACAGAGAACATGAACCAAGTTGTATGGATTGAGGGAGAAGTCAAGTCGAAATACACATCTCTGATGTATCAGGTAATTCTCTGAGGTA
E T E N M N Q L Y G F E G E V K S K Y T S L M Y Q V F S E V

```


ACACCACCCAACAATAATGCAAAGAAGAAGAGTGAGAAGTGTGAATAGTTGACTAATTCCACTGTCTGTCATGAAGTTAATAACTTTCC
T P P N N N A K K K K .

TTCTGTTCTGTTGTCTATATATAATATCTCTGTTGTGCAATGTAGTTTTCATGTACCACA

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGAGGGTACTACTCTGTGGAGACAGTCACTCTTTTAGTAGCTCTTAAGGTTCTGTTTCCCCAAACGAATCACCATCCTGCGAGGCAACCAC
R G Y Y S V E T V T L L V A L K V R F P K R I T I L R G N H 90
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAAAGCCGACAGATCACACAAGTATATGGGTTTTACGACGAATGTCTACGGAAGTACGGAATGCAAATGTTTGAAATACTTCACAGAT
E S R Q I T Q V Y G F Y D E C L R K Y G N A N V W K Y F T D 180
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTGTTTGACTTCCTTCCGCTGACGGCACTGGTGAAGGACAGATTTTCTGCCTACATGGAGGATTGTCCCCCTCCATCGACACGCTGGAC
L F D F L P L T A L V E G Q I F C L H G G L S P S I D T L D 270
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CATATCCGTGCCCTGGACCGTCTACAGGAGGTTCCCCATGAAGGTCGGATGTGTGACCTGCTGTGGTCAGACCCTGATGACAGGGGAGGC
H I R A L D R L Q E V P H E G P M C D L L W S D P D D R G G 360
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGGGGGATCTCCCCCGGGGGCAGGCTACACCTTTGGACAGGACATCTCCGAGACCTTTAACCACAGCAATGGGGCTCACCTGGTGTCC
W G I S P R G A G Y T F G Q D I S E T F N H S N G L T L V S 450
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGAGCACATCAGCTGGTCATGGAGGGCTACAACCTGGTGCCACGATCGGAACGTGGTGACCATCTTCAGCGCACCTAATTACTGCTACCGC
R A H Q L V M E G Y N W C H D R N V V T I F S A P N Y C Y R 540
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGTGGCAACCAGGCCGCCATCATGGAACCTTGATGACACTCTGAAGTACTCCTTCTTGACAGTTTGACCCCGCTCCACGCCGCGGCGAGCCA
C G N Q A A I M E L D D T L K Y S F L Q F D P A P R R G E P 630
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGTCCAGGAGACGCACCCCCGACTACTTCCTGTAAACAGTGAAGGTTTTCGACTCCTGACTGACCCCTCTTTAAAAAAGTGGACCATGCC
H V T R R T P D Y F L . 720

CTGGCCATATACTAAACCACAAAAAAGGGACAAAATTCAGCCAGCAGAAACATTTAAACGTGTCTTTCTGTTGTTAACCAGATATGT

GTGTGGTGGTATCAATTTGTTTAAATTGTAACAAAAAGTTAGGTCTGTACGACAATACGGGGATTTTGATCTTTACTCTA
L₁

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----90
GAAACAGTCGCTGGAGACCATCTGCTTGCTGTAGCGTACAAGATCAAGTACCCGGAGAACTTCTTCTCTACTCAGAGGCCAACACGAGTG
K Q S L E T I C L L L A Y K I K Y P E N F F L L R G N H E C
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----180
TGCGAGCATCAATCGCATCTACGGATTTTACGATGAGTGTAAAAGAAGGTACACAATAAAGCTATGGAAGACGTTTACGGACTGTTTCAA
A S I N R I Y G F Y D E C K R R Y T I K L W K T F T D C F N
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----270
CTGCCTGCCCATCGCAGCCATCATAGACGAGAAGATCTTCTGCTGTCTACGGAGGACTGTCTGCCTGACTTGCAATCCATGGAACAGATCCG
C L P I A A I I D E K I F C C H G G L S P D L Q S M E Q I R


```

-----+-----+-----+-----+-----+-----+-----+-----+
GCGCATCATGCGGCCGACAGACGTCCCGGACCAGGGCCTGCTGTGCGACCTGTTGTGGTTCGGACCCCGACAAGGACGTGATGGGCTGGGG
  R I M R P T D V P D Q G L L C D L L W S D P D K D V M G W G
-----+-----+-----+-----+-----+-----+-----+-----+
450
CGAGAACGACCGCGCGTTCCTTCACGTTTCGGCCCCGAGGTGGTGGCCAAGTTCCTACACAAACATGACCTGGACCTCATATGCAGAGC
  E N D R G V S F T F G P E V V A K F L H K H D L D L I C R A
-----+-----+-----+-----+-----+-----+-----+-----+
540
ACATCAGGTGGTGAAGACGGGTACGAGTTCTTCGCCAAGCGACAGCTGGTCACGTTGTTCTCAGCCCCCACTACTGCGGGGAGTTTGA
  H Q V V E D G Y E F F A K R Q L V T L F S A P N Y C G E F D
-----+-----+-----+-----+-----+-----+-----+-----+
630
CAAGCGAGGGGCAATGATGAGCGTAGACGAGACCCTCATGTGTTCTTTCAGATCTTGAAACCTGCGGACAAGAAGAAGTTCCCGTACGG
  N A G A M M S V D E T L M C S F Q I L K P A D K K K F P Y G
-----+-----+-----+-----+-----+-----+-----+-----+
720
TGGCTTGAACGCGAGGCCGCGCTGTACGCGCCCCCGCGGCCAGAACAAAAACAAAAGGGCAAAGCATAGACCCCCCTCTCCTGTCCCTT
  G L N A G R P V T P P R G Q N K N K K G K A .
-----+-----+-----+-----+-----+-----+-----+-----+
810
TGGCAATATGTAGACAACTGCAGTGTAATGTAATGTAAAGTAAGGTTTGGGGAGAGAATCATCAGAGAAAGGCAGGCTGAAATTCTGTTG
-----+-----+-----+-----+-----+-----+-----+-----+
900
CAGTTGGCCTCTCTAGTAGGTATTCCAGGTAGATGGGTATTG

```

amphioxus amPP2B

```

-----+-----+-----+-----+-----+-----+-----+-----+
GGTACTTCAGCATAGAGTGGCTTCTTTACCTGTGGGCTCTGAAGATGCTCTACCCCAACAACACTATACCTGTTGAGAGGCAACCACGAGT
  Y F S I E C V L Y L W A L K M L Y P T T L Y L L R G N H E
-----+-----+-----+-----+-----+-----+-----+-----+
180
GCAGGCATCTCACAGAGTACTTCACCTTCAAACAAGAGTGTAATAAAGTACAAAGAAGAGGTGTATGATGCCTGTATGGATGCGTTTG
  C R H L T E Y F T F K Q E C K I K Y K E E V Y D A C M D A F
-----+-----+-----+-----+-----+-----+-----+-----+
270
ACTGCCTGCCGCTGGCAGCCCTGATGAACCAACAGTTCTGTGTGTGCACGGCGGCCTGTCCCCGAAATCCACTCACTGGACGACATTA
  D C L P L A A L M N Q Q F L C V H G G L S P E I H S L D D I
-----+-----+-----+-----+-----+-----+-----+-----+
360
CAAAGTTGGACAGATTCAAGGAGCCGCGGCATTTGGCCCCATGTGTGACCTGCTGTGGTTCAGATCCATTAGAAGACTTCGGTAATGAAA
  K L D R F K E P P A F G P M C D L L W S D P L E D F G N E
-----+-----+-----+-----+-----+-----+-----+-----+
450
AGAACGCAGAACACTTCAGTCATAACAGTGTTAGAGGATGCTCCTACTTCTATAGCTATGCGGCAACTTGTGAGTTCCTGCAACAGAACA
  K N A E H F S H N S V R G C S Y F Y S Y A A T C E F L Q Q N
-----+-----+-----+-----+-----+-----+-----+-----+
540
ACCTGTTGTCAATAATTTCGAGCACATGAAGCGCAAGATGCAGGGTATCGAATGTACAGAAAGAGCCAGGCTACTGGTTTTCTTCACTTA
  N L L S I I R A H E A Q D A G Y R M Y R K S Q A T G F P S L
-----+-----+-----+-----+-----+-----+-----+-----+
630
TTACGATTTTCTCTGCACCCAACCTACCTAGATGTTTACAACAACAAGCTGCTGTATTGAAGTATGAAAACAATGTAATGAATATCCGGC
  I T I F S A P N Y L D V Y N N K A A V L K Y E N N V M N I R
-----+-----+-----+-----+-----+-----+-----+-----+
720
AGTTCAACTGCTCGCCACATCCCTATTGGCTACCCAACCTCATGGATGTCTTCACCTGGTCCCTGCCGTTTGTAGGGGAGAAGGTTACAG
  Q F N C S P H P Y W L P N F M D V F T W S L P F V G E K V T
-----+-----+-----+-----+-----+-----+-----+-----+
810
AGATGCTTGTGAATGTGCTAAACATCTGTTTCAGATGATGAACTCATGACAGAGGGTGACGACACCTTTGAAGCGACAGATTCGGACGATG
  E M L V N V L N I C S D D E L M T E G D D T F E A T D S D D

```

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGCAAGGAGGTTCTGTGACAGAACGGAAGGAGATCATTAGGAACAAGATCCGAGCAATCGGCAAGATGGCCAGAGTCTTCACTGTATTAC
G Q G G S V R E R K E I I R N K I R A I G K M A R V F T V L
-----+-----+-----+-----+-----+-----+-----+-----+-----+
900
GAGAGGAGAGTGAGAGTGTCTCTGCATCTGAAGGGACTGACCCCAACTGGGCAGCTTCCCCTCGGAGCCCTGTCTGGAGGAAAAGACACAC
R E E S E S V L H L K G L T P T G Q L P L G A L S G G K D T
-----+-----+-----+-----+-----+-----+-----+-----+-----+
1080
FACAAAGTGCAATCCAGGGTTATAGTGTCAACACCAAGATCCAGTCCTTCGAGGAAGCCAAGTCGCTCGACCGGGTGAACGAGAGGATGC
L Q S A I Q G Y S V N T K I Q S F E E A K S L D R V N E R M
-----+-----+-----+-----+-----+-----+-----+-----+-----+
1170
CGCCGCGCAAGGACGGCACCATGCCATCCAAAGAGTCCATCGAGGACAAGAAGGATGGAAAGGATTCTGCATAAAAAGGCATTGCTGGCG
P P R K D G T M P S K E S I E D K K D G K D S A .
-----+-----+-----+-----+-----+-----+-----+-----+-----+
1260
CATTAGTATTGTTTCATGGGGCCAGCAAAGGGTTCATCCTCTGAAGGCATGGGGCTTCCTTCCCCCTCCCATTACATACTGTACAGTGTA
-----+-----+-----+-----+-----+-----+-----+-----+-----+
1350
TTGTATAACAGCCATAACAGAAATGATTATGTAACCTGTGTGTTTATAGTCCTAATCTCAGATCTGTTTCTGAAGATTGCTCGGGTATC
-----+-----+-----+-----+-----+-----+-----+-----+-----+
1440
CAGTACTAAAATGCAGTTACTTACATAAACCATGAGGGATAAACGCAATGACAACATTGTTCTCTACCATTGTTTGCAAGCAAACCTGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+
1530
AAAGGTTACCCAA

```

amphioxus amPP5

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGAGGTTTCATTTTCAGTGGAGTGTATACTGGTGCTGTTGGATGGAACTCTTGATACCAGATAGCTTCTTCATGGCAAGGGGTAACCAT
R G S F S V E C I L V L F G W K L L I P D S F F M A R G N H
-----+-----+-----+-----+-----+-----+-----+-----+-----+
180
GAGACAGAGACCATGAACCAAATGTATGGGTTTCGAAGGAGAGGTCAAGACTAAGTACTCTGCACAGATGGCGGAGCTCTTCACAGAAGTC
E T E T M N Q M Y G F E G E V K T K Y S A Q M A E L F T E V
-----+-----+-----+-----+-----+-----+-----+-----+-----+
270
TACAACCTACTACCGCTAACACACTGTATCAACAGCAAGATCCTGTGTATGCATGGAGGGCTCTTCAGTAGGGACGATGTTACACTCGAG
Y N L L P L T H C I N S K I L C M H G G L F S R D D V T L E
-----+-----+-----+-----+-----+-----+-----+-----+-----+
360
GAAATCAAGAAGATAGACAGGAATAGACAACCTCCAGACTCAGGTGTGATGTGTGAGCTGTTGTGGTCAGATCCTCAAATCCTGCCTGGT
E I K K I D R N R Q P P D S G V M C E L L W S D P Q I L P G
-----+-----+-----+-----+-----+-----+-----+-----+-----+
450
CGAGCCCCTAGCAAGCGTGGTGTAGCCATCCAGTTTGGACCTGATGTGACGCAAAAGTTCCTGGAGCTCAACAACCTAGACTACATCATC
R A P S K R G V A I Q F G P D V T Q K F L E L N N L D Y I I
-----+-----+-----+-----+-----+-----+-----+-----+-----+
540
CGGAGCCATGAAGTCAAACCCCTAGGGTACGAGGAGGCCCCACAACGGAAAGTGTATCACTGTCTTCTCAGCCCCCACTACTGTGACCAA
R S H E V K P L G Y E E A H N G K C I T V F S A P N Y C D Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+
630
ATGGGAAACAAAGGTGCTTTTCATTACGCTCAACCCACCCGACCTATCTCCAAAATTCACAACGTATGAGTGTGTGCCTCACCCAAATGTG
M G N K G A F I T L N P P D L S P K F T T Y E C V P H P N V
-----+-----+-----+-----+-----+-----+-----+-----+-----+
720
AAGCCCATGGCATACGCCAACTCACTGCTGGCATCCTTCATGTGATGATTGCCCAAGCTATAGTTCTTTTCATGTGTTAGTTACAGTTTTG
K P M A Y A N S L L A S F M .

```

CTTTCCTGTTATGGCTTTTCTACTTTTGTTTATAATAATATACAACATTTTACAATGTGTCAGCATTGAAATAAAAAGATTGCATTTG

PTCATGGTCACTCGAAGTAATTCTGGTCCTGTTTGGGCTTTAAAGTGCTTTACCCTAACCACTTCTACCTTCTCCGTGGAAACCACGAGAC
S W S L E V I L V L F G F K V L Y P N H F Y L L R G N H E T 180

-----+-----+-----+-----+-----+-----+-----+-----+
TGACAACATGAACCAAATGTATGGGTTTGAGGGAGAAGTGAAGGCCAAAGTACACCTCACAGATGTTTGACCTTTTCAGCGAGGTCTTCCA
D N M N Q M Y G F E G E V K A K Y T S Q M F D L F S E V F Q 270

-----+-----+-----+-----+-----+-----+-----+-----+
GTGGCTACCGCTTGCACAGTGTATCAACGGCCGCGTGCTGGTAATGCATGGCGGCCTCTTCAGGGAAGATGGTGTTACTCTGGACAGCAT
W L P L A Q C I N G R V L V M H G G L F R E D G V T L D S I 360

-----+-----+-----+-----+-----+-----+-----+-----+
CCGAAACATTAATCGGAACAGACAACCACTGACTCTGGGCCGATGTGTGACCTTCTGTGGTCAGACCCTCAACCAGAGAACGGAAGGTC
R N I N R N R Q P P D S G P M C D L L W S D P Q P E N G R S 450

-----+-----+-----+-----+-----+-----+-----+-----+
GTGAGTAAGCGTGGAGTGAGCTGTCAGTTTGGGCCGGACGTGACCCATCATTCTCTCGAGGAGAACGGCCTTGACTTGATCATCCGTAG
V S K R G V S C Q F G P D V T H H F L E E N G L D L I I R S 540

-----+-----+-----+-----+-----+-----+-----+-----+
CCACGAGGTGAAGCCTGAGGGTTATGAGGTGGCGCATGGTGGTTCGCTGCATTACAGTCTTCTCTGCTCCAACTACTGTGACCAAATGGG
H E V K P E G Y E V A H G G R C I T V F S A P N Y C D Q M G 630

-----+-----+-----+-----+-----+-----+-----+-----+
AAACAAAGGAGCTTACATCCACCTCAAAGCCCCTGAATTAAAACCGACGTTCCAACAGTTTACTGCTGTGGAACATCCTAATATGAGACC
N K G A Y I H L K A P E L K P T F Q Q F T A V E H P N M R P 720

-----+-----+-----+-----+-----+-----+-----+-----+
GATGGCTTACGCTAATGCACTAATGCAGTTGGGTATCTTGTGATGAAGCCGGTGCATACCTGAGCAGCGACAAATGACGTCGTGTTTTTT
M A Y A N A L M Q L G I L . 810

-----+-----+-----+-----+-----+-----+-----+-----+
TTTTT

CTATTTCAGCATTGAGTGTGTGCTGTACTTGTGGGCGCTAAAAAATACCTGTATCCTACCACACTTTTCTGTGTGCGTGGCAACCACGAGTG
Y F S I E C V L Y L W A L K I L Y P T T L F L L R G N H E C 180
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCGACATCTCACAGAATACTTCACCTTCAAGCAGGAATGCAAGATAAAAATACAGAGACAGTGTACGATGCCTGCATGGAAGCGTTTCGA
R H L T E Y F T F K Q E C K I K Y T E H V Y D A C M E A F D 270
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTGCCTGCCCTTGCTGCCCTCATGAATCAGCAGTTTCTGTGCGTGCATGGTGGACTTTCTCCCGAGGTCCACAATCTGGATGATATCCG
C L P L A A L M N Q Q F L C V H G G L S P E V H N L D D I R 360
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAAGCTAGACAGATTTAAAGAACCGCTGCTTTTGGGCCCATGTGTGACTTGCTCTGGTCTGATCCATCGGAGGATTTTGGCAGTGAAAA
K L D R F K E P P A F G P M C D L L W S D P S E D F G S E K 450
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AACTCAAGAATACTTCTCTCACAAACACAGTTCGAGGTGCTCATATTTCTACAGCTACGCAGCTGTTTGCGACTTTTGCAGAACAA
T Q E Y F S H N T V R G C S Y F Y S Y A A V C D F L Q N N N

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTTTTCTGAATTATACGAGCACATGAAGCACAAGATGCAGGTTATAGAATGTATAGGAAAAGCCAAACTACGGGATTCCCATCCCTAAT
  L L S I I R A H E A Q D A G Y R M Y R K S Q T T G F P S L I
-----+-----+-----+-----+-----+-----+-----+-----+-----+
630
CACGATCTTCTCTGCACCAAACACTACTTGGATGTTTACAACAACAAAGCTGCAGTATTAAAGTATGAGAATAATGTGATGAACATCCGGCA
  T I F S A P N Y L D V Y N N K A A V L K Y E N N V M N I R Q
-----+-----+-----+-----+-----+-----+-----+-----+-----+
720
GTTTAACTGCTCTCCTCACCCCTTATTGGCTGCCCAATTTTCATGGATGTGTTTACATGGTCTCTGCCCTTTGTTGGTGAAAAAGTGACGGA
  F N C S P H P Y W L P N F M D V F T W S L P F V G E K V T E
-----+-----+-----+-----+-----+-----+-----+-----+-----+
810
GATGTTGGTCAATGTGCTGAACATTTGCTCCGACGATGAACTGGTTACAGATGCAGACGAGTCTCTCGACGGCCCGACAACCTGCTGCTCG
  M L V N V L N I C S D D E L V T D A D E S L D G P T T A A R
-----+-----+-----+-----+-----+-----+-----+-----+-----+
900
AAAGGAAGTCATTTCGGAACAAAATAAGAGCCATTGGGAAAATGGCTAAGATGTTTTCTGTTCTCAGAGAGGAGAGTGAGAGTGTGCTGAC
  K E V I R N K I R A I G K M A K M F S V L R E E S E S V L T
-----+-----+-----+-----+-----+-----+-----+-----+-----+
990
CCTAAAAGGTCTTACTCCAACCTGGCATGTTGCCTTCTGGGGTGCTATCAGGAGGACGTCAGACCCTACAAAGCGCGGCTTTTGAGGCAGT
  L K G L T P T G M L P S G V L S G G R Q T L Q S A A F E A V
-----+-----+-----+-----+-----+-----+-----+-----+-----+
1080
TGAGGCAAAACAAAGCCATCAGGGGCTTCTCGCCCCAACAAAAGATCACAAAGTTTTGAGGAAGCAAAGGGGCTGGATCGCATCAACGAACG
  E A N K A I R G F S P Q Q K I T S F E E A K G L D R I N E R
-----+-----+-----+-----+-----+-----+-----+-----+-----+
1170
CATGCCGCCACGAAGGGACACCATTTCACACGATGGTTCTATGAGTCCCATCAACAATTCGGCTGTGGGTGATAATGGCACTGGCAGGAA
  M P P R R D T I S N D G S M S P I N N S A V G D N G T G R N
-----+-----+-----+-----+-----+-----+-----+-----+-----+
1260
TGTACAATGGAAGCAGGCATGGTTTGACACTGATTTGCAAAGTGAGAAAACGGATGCTTTTTGCTTTGCTTTTCTAGCAGGTACCGAATC
  V Q W K Q A W F D T D L Q S E K T D A F C F A F L A G T E S
-----+-----+-----+-----+-----+-----+-----+-----+-----+
1350
GGCTCGCAAAGTTATGATGCAACTGCAAAAAAGGCAATTCCTCGGGTCCCATGACGAATGAACCACATATAAGTGCATGATGCCTATTTTC
  A R K V M M Q L Q K R Q F L G S H D E .
-----+-----+-----+-----+-----+-----+-----+-----+-----+
1440
CTGCTTACCAGATACTGTACAA

```

hagfish hgPP2A

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTATTACTCGGTGGAGACAGTCACAAAGCTTGTTGCTTTGAAGGTTAGGTACCGTGAGCGCATTACCATCCTTCGAGGAAACCACGAGAG
  Y Y S V E T V T K L V A L K V R Y R E R I T I L R G N H E S
-----+-----+-----+-----+-----+-----+-----+-----+-----+
180
CCGCCAGATCACGCAAGTGTACGGGTTTTATGACGAGTGTCTGCGCAAATACGGCAATGCCAACGTCTGGAAATATTTTCACGGATCTCTT
  R Q I T Q V Y G F Y D E C L R K Y G N A N V W K Y F T D L F
-----+-----+-----+-----+-----+-----+-----+-----+-----+
270
CGATTACCTGCCCTCACGGCACTTGTTGATGGCCAGATATTTTGCCTTCACGGTGGGCTCTCGCCATCCATAGACACACTGGATCACAT
  D Y L P L T A L V D G Q I F C L H G G L S P S I D T L D H I
-----+-----+-----+-----+-----+-----+-----+-----+-----+
360
CCGAGCTCTGGATCGCCTGCAAGAGGTTCCACACGAGGGCCCCAATGTGTGACCTGCTGTGGTCAGACCCCGATGATCGTGGTGGATGGGG
  R A L D R L Q E V P H E G P M C D L L W S D P D D R G G W G
-----+-----+-----+-----+-----+-----+-----+-----+-----+
450
GATCTCACCGCGTGGTGCCGGTTACACATTTGGCCAGGATATATCTGAGACCTTCAATCACAGCAATGGGCTTACACTGGTTTCTCGTGC
  I S P R G A G Y T F G Q D I S E T F N H S N G L T L V S R A

```

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACATCAGCTGGTTATGGAGGGATACAATTGGTGCCACGATCGTAATGTGGTGACTATTTTCAGTGCTCCAACTACTGCTACCGTTGTGG
H Q L V M E G Y N W C H D R N V V T I F S A P N Y C Y R C G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAACCAGGUTTCCATTATGGAGTTGGATGACACTCTGAAATACTCCTTTTTGCAGTTTGATCCTGCGCCACGTCGTGGAGAGCCTCACGT
N Q A S I M E L D D T L K Y S F L Q F D P A P R R G E P H V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CACTCGGCGTACTCCAGACTACTTTCTGTAGGGGTGACTACCTGGAAAAGAATAAAATAATATAATGGTGACCTACGACCTCGTCAACC
T R R T P D Y F L .

```

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TATTATCGTCCGAGATTGTTGTTTCCTTCATACCAAGCTTTGACTCCTATATGAGCCCTTTAAACACCCGTTTCCATTCCCAACCCCTTC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCCTTCCCTAATGGATCCCCAAGGTTTTAAAGTGCACCATATGGCTACAAAATGAAACGTTTGCCTTGGATTATCAAGATTCACACTGCA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGCAAGTTTTGCTTTGATCCGCCATTGGTCGTGCAATTATTGATTCACTTGGAAAAATGTTATTTTGCATTAACAATCTAGCATTCATA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AG

```

hagfish hgPPlb

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAAGCAGTCCCTAGAGACAATTTGCTTGCTTCTGGCCTACAAGATCAAGTATCCCGAGAATTTCTTTTGGCTGCGTGGAATCACGAGTG
K Q S L E T I C L L L A Y K I K Y P E N F F L L R G N H E C
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGCTTCTATCAACCGCATCTATGGCTTCTATGATGAATGCAAACGACGCTTCAACATCAAGCTTTGGAAAACATTTACAGACTGCTTCAA
A S I N R I Y G F Y D E C K R R F N I K L W K T F T D C F N
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTGCCTTCCCATTCAGCCATCGTTGATGAGAAAATATTCTGCTGCCATGGAGGCCTGTCTCCCGACCTGCAGTCGATGGAACAGATTTCG
C L P I A A I V D E K I F C C H G G L S P D L Q S M E Q I R
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
LCGCATAATGCGACCCACAGATGTTCTGATACAGGCCTGCTATGTGATTTGCTCTGGTCGGACCCAGACAAAGATGTGCAGGGATGGGG
R I M R P T D V P D T G L L C D L L W S D P D K D V Q G W G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGAAAATGATCGTGGAGTCTCCTTCACTTTTGGCGCTGATGTTGTCAGCAAGTTCCTGAACCGGCATGACTTGGACCTAATTTGCCGTGC
E N D R G V S F T F G A D V V S K F L N R H D L D L I C R A
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCACCAGGTGGTAGAGGATGGATACGAGTTTTTTGCCAAACGACAGCTTGTAAACCCTGTTCTCGGCTCCGAACTACTGTGGAGAGTTTGA
H Q V V E D G Y E F F A K R Q L V T L F S A P N Y C G E F D
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAACGCAGGAGGCATGATGAGTGTTGATGAAACCCTCATGTGCTCCTTCCAGATCTTGAAGCCTTCGGAGAAGAAAGCAAAGTACCAGTA
N A G G M M S V D E T L M C S F Q I L K P S E K K A K Y Q Y
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGTGGCCTAAACTCGGGCCGCCCTGTGACCCACACGCTCAGCCGTACCACCACCGAAGAAGCGATAAAAATGATTTAATAGCCTCGG
G G L N S G R P V T P P R S A V P P P K K R .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AACCGAATGGTTATCACTTGTAGATCTTCACCAGACTTATTAGATGAGATATAGGGGTGTTGCACATAAATAGTCACCAGGTACAGTAAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAATTCCTGGTGGTCCCAACATCAAGGATTGGTCTCGTGTATTTTTT

```


[illegible]

```

-----+-----+-----+-----+-----+-----+-----+-----+
GTTCAACTGTTCCCCACATCCCTACTGGTTGCCCAACTTCATGGATGTCTTCACCTGGTCCTTGCCATTTGTTGGGGAGAAAAGTGACTGA
  F N C S P H P Y W L P N F M D V F T W S L P F V G E K V T E
-----+-----+-----+-----+-----+-----+-----+-----+
310
GATGTTGGTGAATGTGCTCAACATCTGTTCTGATGATGAAGTCAATTCAGAAAGGAGACGATGGCTTTGATGGTGCCTCTGCTGCTGCCCCG
  M L V N V L N I C S D D E L I S E G D D G F D G A S A A A R
-----+-----+-----+-----+-----+-----+-----+-----+
900
TAAAGAGGTGATTAGGAACAAGATTGAGCAATTGGCAAGATGGCCAAAATGTTTTCTGTCCTCCGAGAGGAGAGTGAAAGTGTGCTGAC
  K E V I R N K I R A I G K M A K M F S V L R E E S E S V L T
-----+-----+-----+-----+-----+-----+-----+-----+
990
ACTAAAGGGCCTGACCCCTACAGGAATGCTTCCCAGTGGAGTGTGGCTGGAGGGAAACAAACCCTTCAAAGCGGACTGTTGAGGCTAT
  L K G L T P T G M L P S G V L A G G K Q T L Q S A T V E A I
-----+-----+-----+-----+-----+-----+-----+-----+
1080
TGAGGCTGATGAAGCCATCCGGGGATTCTCACCACAGCATAAGATCAGCAGTTTTGCTGAAGCCAAGGGTTGGATCGAATCAACGAAAG
  E A D E A I R G F S P Q H K I S S F A E A K G L D R I N E R
-----+-----+-----+-----+-----+-----+-----+-----+
1170
GATGCCTCCACGACGTGATGGAATCGCTACTGATGGCAGCTTTAACTCCATTAACAAGGCAGTCTCTTCTGAGACGAACGGCACTGAGAG
  M P P R R D G I A T D G S F N S I N K A V S S E T N G T E S
-----+-----+-----+-----+-----+-----+-----+-----+
1260
TGGTAACATTCAATGATCCCTTTGCAGTTTCCAGTCCAAGGCATGATTGTATCATTGTCCAAGCGGTTTGACACTCTTGCCCTCTGATT
  G N I Q .
-----+-----+-----+-----+-----+-----+-----+-----+
1350
ACTGCTGCATGTTCTGGGGAGCCAGATGTTGGTGACAGATTCCAGTAACAGAGTGATAGAAATTGTATTGTGATTGTTGAAAGTTAAGTT
-----+-----+-----+-----+-----+-----+-----+-----+
1440
ACCTCAGATGGGATATGCCTTTGATGCTTCTAGACAAGCTGGTTAAATATTACACA

```

ray ryPP2Bb

```

-----+-----+-----+-----+-----+-----+-----+-----+
CTACTTCAGTATCGAGTGTGTTTTGTATCTGTGGTCGTTGAAAATTCTGTATCCAAAGACACTGTTTTTACTCCGTGGAACCACGAATG
  Y F S I E C V L Y L W S L K I L Y P K T L F L L R G N H E C
-----+-----+-----+-----+-----+-----+-----+-----+
180
TAGACATTTGACAGAATATTTCACTTTCAAACAAGAATGTAAAATCAAGTATTCAGAACAGATTTATGATGCATGTATGGAAGCCTTCGA
  R H L T E Y F T F K Q E C K I K Y S E Q I Y D A C M E A F D
-----+-----+-----+-----+-----+-----+-----+-----+
270
CTGCCTACCCCTTGCGGCACTGATGAATCAGCAATTTCTTTGTGTACACGGAGGCCTTTCCCCTGAAATACACACCCTGGATGATATAAA
  C L P L A A L M N Q Q F L C V H G G L S P E I H T L D D I K
-----+-----+-----+-----+-----+-----+-----+-----+
360
AAAGTTAGACAGATTTAAAGAGCCACCAGCATTTGGTCCTATGTGTGATCTGTTATGGTCAGATCCATTGGAAGATTTTGGAATGAAAA
  K L D R F K E P P A F G P M C D L L W S D P L E D F G N E K
-----+-----+-----+-----+-----+-----+-----+-----+
450
GACACAGGAATATTTTCAGTCATAATACAGTGAGAGGGTGCTCTTATTTTTATAGTTATCCTGCCGTTTGTGAGTTCCTGCAAAACAATAA
  T Q E Y F S H N T V R G C S Y F Y S Y P A V C E F L Q N N N
-----+-----+-----+-----+-----+-----+-----+-----+
540
TTTATTATCAGTTATTCGAGCCCATGAAGCACAAAGATGCAGGCTATCGCATGTACAGGAAGAGCCAAACTACGGGTTTCCCATCACTTAT
  L L S V I R A H E A Q D A G Y R M Y R K S Q T T G F P S L I
-----+-----+-----+-----+-----+-----+-----+-----+
630
CACAAATCTTTTCTGCACCAAACTACCTTGATGTTTACAATAACAAAGCTGCAGTATTGAAATATGAAAACAATGTTATGAATATTCGGCA
  T I F S A P N Y L D V Y N N K A A V L K Y E N N V M N I R Q

```

```

-----+-----+-----+-----+-----+-----+-----+-----+
ATTCAATGTTCTCTCTCATCCTTACTGGCTACCAAATTTTATGGATGTTTTTACTTGGTCTTTACCATTTGTTGGTGAAAAAGTGACAGA
  E N C S P H P Y W L P N F M D V F T W S L P F V G E K V T E
-----+-----+-----+-----+-----+-----+-----+-----+
GATGTTGGTGAATGTTCTAAATATCTGCTCTGATGATGAACATCATGACAGATGGAGATGAGATATTGATGCAAATGCTGCTGCAGCTCG
  M L V N V L N I C S D D E L M T D G D E I F D A N A A A A R
-----+-----+-----+-----+-----+-----+-----+-----+
AAAAGAGGTTATCAGGAACAAAATTCGAGCGATTGGCAAAATGGCAAGAGTGTTTTCTGTGCTCAGAGAGGAGAGTGAGAGTGTTTTAAC
  K E V I R N K I R A I G K M A R V F S V L R E E S E S V L T
-----+-----+-----+-----+-----+-----+-----+-----+
TTTGAAGGTTCTGACTCCAACCGGTATGCTGCCTAGTGGAGTGCTGTCAGGAGGCAAACAGACTTTACAAAGCGCTACTGTTGAGGCTAT
  L K G L T P T G M L P S G V L S G G K Q T L Q S A T V E A I
-----+-----+-----+-----+-----+-----+-----+-----+
TGAGGCCATTGAGGGTGTGAAACTCATGAAGCCATTGAGGATTTACACCTCAGCATAAAATTAGTAGTTTTGAAGAAGCAAAGGCCT
  E A I E G V E T H E A I R G F T P Q H K I S S F E E A K G L
-----+-----+-----+-----+-----+-----+-----+-----+
GGACCGTATCAATGAGAGAATGCCGCCTCGAAGAGATGCAGTATCCAATTCTGGCATTGTTTCACTCAACTGGGAAAATGAATATGTCCGA
  D R I N E R M P P R R D A V S N S G I V H S T G K M N M S E
-----+-----+-----+-----+-----+-----+-----+-----+
ATCAAAATGGGACAGATGACAACAGCAACATACAGTGATCTGCTTTTAAGACTTCAATATCTCTGCAGTTGTACTGTGTCAACTGTTTGGT
  S N G T D D N S N I Q .
-----+-----+-----+-----+-----+-----+-----+-----+
GTTCCAGCCTCTGATATCCAAGACAGCAGCTGCTCATTGAGAAGCAAAGAGTTTGAATAAATTTATAAAAAGAAAACACAAACCCCAA
  AACTTCCACTCAGAAAATTAGTGTTAGAAAAGCTGGAAACCTTA

```

ray ryPP1b

```

-----+-----+-----+-----+-----+-----+-----+-----+
CGGGGGAAGCAATCTCTGGAGACCATTGCGCTTCTCCTAGCATACAAGATCAAATACCCAGAAAACCTTCTTCCTTTTGCGGGGAAATCAT
  R G K Q S L E T I C L L L A Y K I K Y P E N F F L L R G N H
-----+-----+-----+-----+-----+-----+-----+-----+
GAGTGCGCCAGCATCAATCGTATCTATGGGTTCTACGATGAGTGCAAGAGACGTTTTAATATCAAACCTCTGGAAGACATTTACGGACTGC
  E C A S I N R I Y G F Y D E C K R R F N I K L W K T F T D C
-----+-----+-----+-----+-----+-----+-----+-----+
TTCAACTGTCTTCCAATTGCAGCCATTGTGGATGAAAAAATTTTCTGTTGTCATGGGGGGCTGTCAACGATTGCAATCAATGGAACAG
  F N C L P I A A I V D E K I F C C H G G L S P D L Q S M E Q
-----+-----+-----+-----+-----+-----+-----+-----+
ATTTCGAGAATCATGAGACCAACAGATGTTTCTGACACAGGATTGCTTTTGATTTGCTGTGGTCAGATCCTGACAAGGACGTTCAAGGT
  I R R I M R P T D V P D T G L L C D L L W S D P D K D V Q G
-----+-----+-----+-----+-----+-----+-----+-----+
TGGGGTGAGAATGATCGTGGCGTCTCTTTCACCTTTGGAGCTGATGTAGTCAGCAAATTTCTGAACCGGCATGATTGGATTAAATCTGT
  W G E N D R G V S F T F G A D V V S K F L N R H D L D L I C
-----+-----+-----+-----+-----+-----+-----+-----+
CGTGCCCATCAGGTGTTGAAGATGGTTACGAATTTCTTGCTAAACGCCAGCTGGTTACTTTGTTTTCTGCTCCAACTACTGTGGAGAG
  R A H Q V V E D G Y E F F A K R Q L V T L F S A P N Y C G E
-----+-----+-----+-----+-----+-----+-----+-----+
TTTGATAATGCCGGAGGAATGATGAGTGTTGACGAAACATTAATGTGTTTCCTTTTCTGATTCTAAAACCTTCTGAAAAGAAAGCGAAATAC
  F D N A G G M M S V D E T L M C S F Q I L K P S E K K A K Y

```

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGTATGGTGGTCTGAACTCTGGTCGCCCTGTTACGCCCCCTCGCACGGCAAACCTCCCAAGAAGAGGTGAATAGCAAGGACAGCACCA
Q Y G G L N S G R P V T P P R T A N P P K K R .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AACACCTTATCATTGTACACAGAGAAAGCTACGAGTATACCAGTTCATATACTAATGTGTGTAACCTGTAGGCACTGTTCAACCATCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTACTGTCTGTCTTTGTGAGCCTCAATGCTTGTA AAAAGAGAAGTCTTTTGA AAAAATGTTTTCCAGTAAATGCGCAGGTGTGGCATAT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGTAATGTAAATCAATGTAGTACCAGTAACCAGTGCGTATTGGTAATACTTTGGGCTGGGATGAAGGTTAACAAATCTAAGACTGGGATA
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AAA

```

ray ryPPla

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAAGCAGTCATTGGAAACTATCTGTCTTCTGTTGGCTTACAAAATTAATATCCAGAAAATTTTTCTTCTGCGTGGAACCATGAATG
K Q S L E T I C L L L A Y K I K Y P E N F F L L R G N H E C
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGCCAGTATAAACAGAATTTATGGATTTTATGATGAATGTAAAAGGAGGTACAATATTAATTTGTGGAAGACCTTCACAGACTGTTTCAA
A S I N R I Y G F Y D E C K R R Y N I K L W K T F T D C F N
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTGCTTACCAATAGCTGCCATTGTGGATGAAAAGATATTTTGTGTCATGGTGGTTTGTACCGGATCTTCAGTCCATGGAACAAATCAG
C L P I A A I V D E K I F C C H G G L S P D L Q S M E Q I R
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACGAATCATGCGACCCACTGATGTGCCTGATCAAGGACTTCTTTGTGACCTTCTGTGGTCTGATCCAGACAAAGATGTATTGGGATGGGG
R I M R P T D V P D Q G L L C D L L W S D P D K D V L G W G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGAAAATGATCGAGGTGTCTCATTTACATTTGGGGCAGAAGTAGTTGCAAAGTTCCTTCATAAACATGATCTGGATCTGATATGTAGAGC
E N D R G V S F T F G A E V V A K F L H K H D L D L I C R A
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ACACCAGGTTGTAGAAGATGGGTATGAGTTCTTTGCAAAGAGACAATTGGTGACTTTATCTCTGCACCCAATTATTGTGGAGAGTTTGA
H Q V V E D G Y E F F A K R Q L V T L F S A P N Y C G E F D
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TAATGCAGGTGCTATGATGAGTGTAGATGAAACCTTGATGTGCTCTTTTCAGATTTTGAAACCAGCAGAGAAAAAGAAACCCAATACAAG
N A G A M M S V D E T L M C S F Q I L K P A E K K K P N T S
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCGGCCTGTAAACACCACCAAGGAGCATGATCAGAAAACAAGCCAAGAAGTAAACCAGGAATTCTGGTGTTGCCTGTGCGACATGTATTA
R P V T P P R S M I T K Q A K K .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TATCAGATTGTCTCCATTTTAACTTTTTTTGTGTAAAGTGACCTGTCAGCTTGCTTGGGTTAGATTACTACTTGTGACTTGTGGTT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TATCTGCTGTGATGAGCTGAATCATATTTGCTGCTGCAGAGCAACAAACATGAGTCAGCCATCTGAGATGGAAGATTATTA AAAATACAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TAGCTCTC

```

ray ryPP4a

90

STTCTACAGTGTAGAAACGTTCCCTCCTGCTCCTCGCTCTAAAGGTGCGATACCCAGACAGGATTACACTGATCCGGGGTAACCATGAGAG

F Y S V E T F L L L L A L K V R Y P D R I T L I R G N H E S

180

TCGTGAGATCACCCAGGTGTACGGCTTCTATGACGAGTGTCTTCGCAAGTACGGCTCTGTGACCGTCTGGAGATACTGTACTGAAATCTT

R Q I T Q V Y G F Y D E C L R K Y G S V T V W R Y C T E I F

270

CGACTATCTCAGTCTGTCAGCCATCATTGACGGCAAGATATTCTGTGTACACGGGGGCTCTCACCCCTCCATCCAGACGCTGGACCAGAT

D Y L S L S A I I D G K I F C V H G G L S P S I Q T L D Q I

360

CCGTACCATCGACCGGAAACAGGAGGTTCCGCACGACGGTCCCATGTGTGACCTGCTCTGGTCTGACCCCGAGGACACGACGGGATGGGG

R T I D R K Q E V P H D G P M C D L L W S D P E D T T G W G

450

TGTGAGCCCACGAGGTGCAGGTTACCTCTTTGGCAGCGACGTGGTGGCACAGTTCAACGCAGCCAACGACATCGACATGATCTGCCGTGC

V S P R G A G Y L F G S D V V A Q F N A A N D I D M I C R A

540

CCACCAGCTCGTCATGGAGGGTTACAAATGGCATTTC AACGAGACTGTGCTTACTGTCTGGTCAGCACCCAACTACTGCTACAGGTGTGG

H Q L V M E G Y K W H F N E T V L T V W S A P N Y C Y R C G

630

GAATGTGGCGGCCATCCTGGAGCTGGACGAGCATCTTCAGAAAGAGTTCATCATCTTCGAGGCAGCGCCTCAGGAGACCCGAGGCATCCC

N V A A I L E L D E H L Q K E F I I F E A A P Q E T R G I P

720

CTCCAAAAACCTGTGGCTGATTACTTCCTTTGAACTGTCCTCCTAGTCCTCGTCGTCCTCCTCCTCCGTCCTCCTCATCGGTCCT

S K K P V A D Y F L .

810

CACCCCTACTCCCTCCCTCGCCCATCCCCTCCAAAAATCACACGTCCAGTTGACCCTGTTTCATCATTTCTCCTCACTATCTTGAGGTCT

900

TACTAT

ray ryPP2Ab

90

CGAGGTTATTATTTCAGTTGAAACAGTCACACTTCTTGTAGCCCTTAAGGTACGTTATCGTGAACGTATCACAATACTTAGAGGGAATCAT

R G Y Y S V E T V T L L V A L K V R Y R E R I T I L R G N H

180

GAAAGCAGACAAATTACACAAGTGTATGGCTTCTATGATGAATGTTTAAGAAAATATGGAAATGCCAATGTGTGGAAATATTTTACTGAT

E S R Q I T Q V Y G F Y D E C L R K Y G N A N V W K Y F T D

270

CTCTTTGATTACCTTCCGCTCACTGCCTTGGTAGATGGTCAGATCTTCTGTCTCCATGGAGGCCTTTCTCCTTCCATTGATACACTGGAT

L F D Y L P L T A L V D G Q I F C L H G G L S P S I D T L D

360

CACATTTCGAGCACTAGATCGCCTTCAGGAGGTTCCACATGAGGGTCCAATGTGTGATTGCTATGGTCGGATCCTGATGATCGTGGGGGC

H I R A L D R L Q E V P H E G P M C D L L W S D P D D R G G

450

TGGGGCATATCCCCCGGGGTGCTGGTTACACCTTTGGTCAGGATATTTCTGAGACTTTTAACCATGCGAATGGCCTTACATTGGTCTCA

W G I S P R G A G Y T F G Q D I S E T F N H A N G L T L V S

540

CGAGCACATCAACTGGTGATGGAGGGTTACAACCTGGTSCCATGATCGCAATGTAGTGACAATTTTCAGTGCTCCAAACTACTGTTATCGT

R A H Q L V M E G Y N W C H D R N V V T I F S A P N Y C Y R


```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
PGTGGGAATCAAGCTGCAATAATGGAACCTGATGACACCCCTTAAATATTCTTTCCTACAGTTTGATCCTGCACCTCGCAGAGGTGAACCA
C G N Q A A I M E L D D T L K Y S F L Q F D P A P R R G E P
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CATGTGACTCGTCGCACCCAGATTATTTCTGTAATAATCTTTTGTACAGTATTGCTATGAATTAAATGTTGGCCTGAAAGAATTGGAAG
H V T R R T P D Y F L .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGCAACAGTAACTTCAGATTATAAGTGAGAATCTCAAACATTTCAGAAACCTGTCATCACGGACCAAAAGACGTGCCATAATAAATCAAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCTCAGTCTGCAAACATCTCATTGTTCTACTTGAGCACCTGTTTCATCAATGCATGCTGATGCGTGAAAGTTATGGGTTGAGACCTGTGGT
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTGAGCCTAGTTTGTAGTTACTTTTGCTTTTTCAGAGATTG

```

ray ryPP2Aa

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTATTATTTCAGTTGAAACTGTGACATTGCTCGTGACACTGAAGGTTTCGATATCGAGAACGCATCACAATATTACGAGGGAACCATGAAAG
Y Y S V E T V T L L V T L K V R Y R E R I T I L R G N H E S
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGACAAATCACACAAGTATATGGGTTTTATGATGAATGTTTAAGAAAATATGGAAATGCTAGTGTGTTGGAAGTACTTCACAGACCTGTT
R Q I T Q V Y G F Y D E C L R K Y G N A S V W K Y F T D L F
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGATTATCTGCCCTTGACAGCCTTAGTAGATGGTCAGATATTTGCCTGCATGGTGGACTTTCTCCATCCATTGACACATTGGATCACAT
D Y L P L T A L V D G Q I F C L H G G L S P S I D T L D H I
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCGAGCCTTGATCGGCTGCAAGAAGTGCCACATGAGGGACCAATGTGTGATCTGCTCTGGTCTGATCCTGATGATCGTGGTGGTTGGGG
R A L D R L Q E V P H E G P M C D L L W S D P D D R G G W G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TATTTCTCCTCGAGGTGCAGGCTATACTTTTGACAGGACATTTCTGAAACTTTCAATCATGCAAATGGGCTCACTCTGGTGTCTCGGGC
I S P R G A G Y T F G Q D I S E T F N H A N G L T L V S R A
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCATCAACTTGTAATGGAGGGATACAACCTGGTGCCATGATCGAAATGTGGTGACAATATTCAGTGCTCCCAACTATTGTTATCGCTGTGG
H Q L V M E G Y N W C H D R N V V T I F S A P N Y C Y R C G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TAATCAAGCTGCTATTATGGAATTAGACGATACACTGAAGTATTCATTTCTTCAGTTTGATCCTGCACCTCGCCGTGGAGAACCTCATGT
N Q A A I M E L D D T L K Y S F L Q F D P A P R R G E P H V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CACTCGACGTACTCCAGACTATTTCTTGTAAGAATTCTTCCGTACACTACAACCACGCCAGCAGTTCATCTGCTTCA
T R R T P D Y F L .

```

ray ryPP5

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGTGGAGGTTATATTAGTCCTCTTCGGATTCAAGCTGCTCTATCCGAACCACTTCCATGTTCTCCGAGGTAACCATGAAACAGATAACAT
V E V I L V L F G F K L L Y P N H F H V L R G N H E T D N M
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GAACCAAATGTATGGCTTTGAAGGTGAGGTGAAGGCCAAGTATACGTCGCAGATGTTTGAGCTCTTCAGTGAGGTGTTCCAGTGGCTCCC
N Q M Y G F E G E V K A K Y T S Q M F E L F S E V F Q W L P

```

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
PCTTGCTCAGTGCATCAACAGTAAAGTTTTGATAATGCACGGAGGGCTCTTCAGTGAAGATGGTATCACCCCTCGAAGATATCAAGAAGAT
  L A Q C I N S K V L I M H G G L F S E D G I T L E D I K K I
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
360
TGATCGAAATCGACAACCTCCAGATTGAGGTCCAATGTGTGACTTGCTCTGGTCAGATCCACAGCCTCAGAATGGGCGATCGGTGAGCAA
  D R N R Q P P D S G P M C D L L W S D P Q P Q N G R S V S K
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
450
GCGTGGTGTGAGCTGCCAGTTTGGGCCAGATGTCACCAAGCGATTCTTGAGGAGAAACAAGCTGGATTACATAATCAGGAGCCACGAGGT
  R G V S C Q F G P D V T K R F L E E N K L D Y I I R S H E V
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
540
GAAGCCAGAGGGTTACGAGGTGACGCACAATGGCAAATGTATCACTGTGTTCTCTGCGCCAAACTATTGTGATCAAAATGGGAAACCAAGG
  K P E G Y E V T H N G K C I T V F S A P N Y C D Q M G N Q G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
630
GGCATACATTACCTCAGGGGCTCGGATCTAAAACCGTTCTTTCATCAATTTCAAACCTGTGCCTCACCCGAATGTTAAACCGATGGCATA
  A Y I H L R G S D L K P F F H Q F Q T V P H P N V K P M A Y
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
720
TGCAAACCTCACTACTGCAGCTGGGAATGATGTAGCATCGGAGCATCTGGTTTGGACAGTAATCCCTCCCACAGGTGGGGAGGGATGGCAT
  A N S L L Q L G M M .
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
810
CAGGGTTTAATTTAATCTTCCTCACTAGATTTGCCATCTCTGAACCAAGTCTTCTACCTTTCTAGCACCACAGCCCTTGATGATGTGACC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
900
TGCCCCATGCCGTGCTAAGTTATGGAGTATAAATCAGCTCTCAAACACGTGGCCCATGCACAGTTTCGTGTTTCAGACATACATCCTTGCCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
990
TTCA

```

ray ryPP4b

```

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CTTCTACAGTGTGAGACATTCTTGCTGCTATTAGCACTGAAGGTTTCGTTATCCTGATAGAATCACACTGATCAGAGGGAACCACGAGAG
  F Y S V E T F L L L L A L K V R Y P D R I T L I R G N H E S
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
180
CAGACAGATTACCCAGGTTTATGGGTTTTATGATGAGTGCTTGCGCAAATATGGATCGATAACTGTGTGGAGATACTGCACTGAGATCTT
  R Q I T Q V Y G F Y D E C L R K Y G S I T V W R Y C T E I F
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
270
TGATTACCTCAGCCTTCTGCTGTTGTGGATGGAAGATCTTCTGTGTGCATGGAGGCCTGTCTCCATCCATCCAGACCCTGGACCAGAT
  D Y L S L S A V V D G K I F C V H G G L S P S I Q T L D Q I
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
360
CAGGACAATTGACCGGAAACAGGAAGTGCCCCACGATGGTCCAATGTGTGATTTGCTGTGGTCTGACCCAGAAGATACAACCTGGCTGGGG
  R T I D R K Q E V P H D G P M C D L L W S D P E D T T G W G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
450
CATTAGCCCCCGTGGAGCTGGCTACCTCTTTGGCAGTGACGTTGTTGCCCACTTCAGCGCTGCCAATAGCATAGACCTGATCTGCCGTGC
  I S P R G A G Y L F G S D V V A H F S A A N S I D L I C R A
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
540
CCACCAGCTGGTGTGAGAGGGTTACAAGTGGCATTTCAATGAAACTGTTCTAACTGTCTGGTCAGCACCCAACTACTGTTACAGGTGTGG
  H Q L V M E G Y K W H F N E T V L T V W S A P N Y C Y R C G
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
630
AAACGTCGCTGCAATTCTTGAACCTGGATGAGCATTTAGAGAAAGAGTTTACTATTTTTGAGGCTGCTCCTCAGGAATCTCGAGGGGTGTC
  N V A A I L E L D E H L E K E F T I F E A A P Q E S R G V S

```

720

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GGCTAAAAAGACAGTTGCTGACTATTTCTCTGAAGGTGTACCATAATATCGTGGAATTACACTCTCAGCACTGGGCCTGCCTAAGGAGG
A K K T V A D Y F L .

810

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GACCAATGTGACAAAATGGTATACTTGTGAACTGCCCTTGCATTCGGAAGACAGACTAGGCCATCTGCCTGCTCCTGGGGAAATGCCATC

900

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TTGCTAGTGCCATGGCAATGGGGCCATGGTGATCTCTTGCAATCCTTGTGGCTCTTCCTTACATTGGGGAAAATAAATAACTGTACGGAA

990

-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
A

付録3 系統樹推定に用いた遺伝子の生物種名、遺伝子名、およびGenBankのアクセッションナンバーを、サブタイプごとに分けて示した。番号は図7の分子系統樹上の番号に対応する。

3-1 図7のPTP遺伝子族分子系統樹推定に用いた遺伝子。番号は図7の分子系統樹上の番号に対応する。

PTPR7

1. *Homo sapiens* STEP (U27831)
2. *Homo sapiens* HPCPTP1 (D64053)
3. *Homo sapiens* LC-PTP (D11327)
4. *Ephydatia fluviatilis* sPTPR7 (AB019131)

PTPN3

5. *Homo sapiens* PTPN1 (M33689)
6. *Gallus gallus* PTF1B (U86410)
7. *Homo sapiens* PTPRF (M25393)
8. *Potamotrygon motoro* ryPTPN3 (AB033582)
9. *Branchiostoma belcheri* amPTPN3 (AB033568)
10. *Eptatretus burgeri* hgPTPN3 (AB033579)
11. *Drosophila melanogaster* Ptp61F (L11253)
12. *Ephydatia fluviatilis* sPTPN3

PTPN4

13. *Homo sapiens* PTP-BAS type 1 (D21209)

PTPN5

14. *Homo sapiens* DKFZp566K0524 (AL050040)

PTPN9

15. *Homo sapiens* DKFZp564F0923 (AL110210)

PTPN8

16. *Homo sapiens* pez (X82676)
17. *Homo sapiens* PTP D1 (X79510)
18. *Ephydatia fluviatilis* sPTPN8 (AB019133)

PTPN7

19. *Homo sapiens* PTP (M64572)
20. *Homo sapiens* PTP (M68941)

PTPN1

21. *Xenopus laevis* PTPX1 (L33098)
22. *Xenopus laevis* PTPX10 (L33099)
23. *Homo sapiens* PTP MEG2 (M83738)
24. *Ephydatia fluviatilis* sPTPN1 (AB019129)

PTPN2

25. *Ephydatia fluviatilis* sPTPN2 (AB019130)
26. *Homo sapiens* 70zpep (AF077031)
27. *Homo sapiens* PTP G1 (D13380)
28. *Homo sapiens* BDP1 (X79568)

PTPR8

29. *Homo sapiens* IA-2/PTP (L18983)
30. *Potamotrygon motoro* ryPTPR8 (AB033585)
31. *Homo sapiens* IAR (AF007555)

PTPN6

32. *Potamotrygon motoro* ryPTPN6b (AB033591)
33. *Homo sapiens* PTP 1D (X70766)
34. *Gallus gallus* Syp (U38620)
35. *Xenopus laevis* SH-PTP2 (U15287)
36. *Eptatretus burgeri* hgPTPN6a (AB033576)
37. *Eptatretus burgeri* hgPTPN6b (AB033578)
38. *Potamotrygon motoro* ryPTPN6c (AB033592)
39. *Homo sapiens* hPTP1C (U15528)
40. *Potamotrygon motoro* ryPTPN6a (AB033580)
41. *Drosophila melanogaster* csw Y1229 (U19909)

42. *Branchiostoma belcheri* amPTPN6 (AB033567)
43. *Ephydatia fluviatilis* sPTPN6 (AB019132)

PTPR5

44. *Potamotrygon motoro* ryPTPR5b (AB033590)
45. *Homo sapiens* PTP γ (L09247)
46. *Gallus gallus* PTP γ (U38349)
47. *Eptatretus burgeri* hgPTPR5a (AB033570)
48. *Potamotrygon motoro* ryPTPR5a (AB033588)
49. *Homo sapiens* PTP ζ (M93426)
50. *Gallus gallus* CPTP ζ (L27625)
51. *Eptatretus burgeri* hgPTPR5b (AB033577)
52. *Branchiostoma belcheri* amPTPR5 (AB033565)
53. *Drosophila melanogaster* Ptp99A (M80539)
54. *Ephydatia fluviatilis* sPTPR5 (AB019127)

PTPR9

55. *Drosophila melanogaster* Ptp69D (M27699)

PTPR1/6

56. *Homo sapiens* LCA (Y00638)
57. *Gallus gallus* ptp lambda (L13285)
58. *Heterodontus francisci* CD45 (U34750)

PTPR4

59. *Potamotrygon motoro* ryPTPR4b (AB033589)
60. *Homo sapiens* PTP α (M34668)
61. *Xenopus laevis* PTP α (U09135)
62. *Gallus gallus* ptp α (Z32749)
63. *Eptatretus burgeri* hgPTPR4 (AB033572)
64. *Homo sapiens* HPTP ϵ (X54134)
65. *Potamotrygon motoro* ryPTPR4a (AB033583)
66. *Branchiostoma belcheri* amPTPR4c (AB033564)
67. *Branchiostoma belcheri* amPTPR4a (AB033562)
68. *Branchiostoma belcheri* amPTPR4b (AB033563)
69. *Branchiostoma belcheri* amPTP10 (AB033561)

PTPR2B

70. *Ephydatia fluviatilis* sPTPR2B (AB019126)
71. *Xenopus laevis* ptprt (AF173857)
72. *Homo sapiens* RPTP-rho (AF043644)
73. *Homo sapiens* hR-PTP μ (X58288)
74. *Eptatretus burgeri* hgPTPR2B (AB033574)
75. *Homo sapiens* h-PTP κ (Z70660)
76. *Potamotrygon motoro* ryPTPR2B (AB033587)
77. *Homo sapiens* hPTP-J (U73727)

PTPR2A

78. *Drosophila melanogaster* Lar (M27700)
79. *Hirudo medicinalis* HmLAR1 (AF017084)
80. *Homo sapiens* PTP δ (L38929)
81. *Potamotrygon motoro* ryPTPR2Ab (AB033584)
82. *Branchiostoma belcheri* amPTPR2A (AB033566)
83. *Eptatretus burgeri* hgPTPR2Aa (AB033569)
84. *Potamotrygon motoro* ryPTPR2Ac (AB033586)
85. *Gallus gallus* CRYP α 1 (L32780)
86. *Homo sapiens* PTP σ (U35234)
87. *Homo sapiens* LAR (Y00815)
88. *Potamotrygon motoro* ryPTPR2Aa (AB033581)
89. *Eptatretus burgeri* hgPTPR2Ab (AB033571)
90. *Hirudo medicinalis* HmLAR2 (AF017083)
91. *Eptatretus burgeri* hgPTPR2Ac (AB033575)
92. *Ephydatia fluviatilis* sPTPR4 (AB019125)

PTPR3

93. *Ephydatia fluviatilis* sPTPR3 (AB019128)
94. *Homo sapiens* DEP-1 (U10886)
95. *Gallus gallus* PTP (AJ238216)
96. *Homo sapiens* HPTP β (X54131)
97. *Eptatretus burgeri* hgPTPR3 (AB033573)
98. *Homo sapiens* PTP-U2 (Z48541)
99. *Gallus gallus* CRYP-2 (U65891)
100. *Rattus norvegicus* PTPRQ (AF063249)
101. *Drosophila melanogaster* Ptp10D (M80465)
102. *Drosophila melanogaster* Ptp4E (L20894)
103. *Branchiostoma belcheri* amPTPR3 (AB033560)
104. *Mus musculus* Esp (U36488)
105. *Homo sapiens* SAP-1 (D15049)

3-2 図27のPP遺伝子族分子系統樹推定に用いた遺伝子。番号は図27の分子系統樹上の番号に対応する。

rdgC

1. *Drosophila melanogaster* rdgC (M89628)
2. *Homo sapiens* PPEF-2 (AF023456)
3. *Homo sapiens* PPP7C (AF027977)
4. *Trypanosoma brucei* AAC14087.1 (AF052832)

PP7

5. *Arabidopsis thaliana* PP7 (AJ000057)

PP5

6. *Saccharomyces cerevisiae* PPT1 (X83099)
7. *Schizosaccharomyces pombe* SPBC3F6.01c (AL022019)
8. *Neurospora crassa* ppt-1 (U89985)
9. *Homo sapiens* PP5 (U25174)
10. *Potamotrygon motoro* ryPP5
11. *Eptatretus burgeri* hgPP5
12. *Branchiostoma belcheri* amPP5
13. *Ephydatia fluviatilis* sPP5
14. *Arabidopsis thaliana* F7D19.19 (AC006931)

PP2B

15. *Homo sapiens* PPP3CA (L14778)
16. *Potamotrygon motoro* ryPP2Ba
17. *Potamotrygon motoro* ryPP2Bb
18. *Eptatretus burgeri* hgPP2B
19. *Homo sapiens* calcineurin A2 (M29551)
20. *Drosophila melanogaster* Pp2B-14D (U30493)
21. *Homo sapiens* calcineurin A (S46622)
22. *Branchiostoma belcheri* amPP2B
23. *Ephydatia fluviatilis* sPP2B
24. *Drosophila melanogaster* CanA1 (M97012)
25. *Neurospora crassa* calmodulin-dependent protein phosphatase (M73032)
26. *Filobasidiella neoformans* CNA1 (AF042082)
27. *Schizosaccharomyces pombe* ppb1+ (D28955)
28. *Saccharomyces cerevisiae* CMP1 (X66490)
29. *Saccharomyces cerevisiae* CMP2 (X54964)
30. *Dictyostelium discoideum* calcineurin (U22397)
31. *Paramecium tetraurelia* protein phosphatase 2B (AF014922)

PP6

32. *Schizosaccharomyces pombe* ppe1+ (D13712)
33. *Saccharomyces cerevisiae* SIT4 (M24395)
34. *Homo sapiens* PP6C (X92972)
35. *Drosophila melanogaster* PPPV6A (X75980)
36. *Arabidopsis thaliana* AAD50050.1 (AC007980)
37. *Dictyostelium discoideum* pppD (AF176121)
38. *Plasmodium falciparum* MAL3P5.5 (AL034556)

PP4

39. *Schizosaccharomyces pombe* SPBC26H8.05c (AL031743)
40. *Dictyostelium discoideum* pppC (AF161253)
41. *Arabidopsis thaliana* PPX-2 (Z22596)
42. *Homo sapiens* PPX (X70218)
43. *Potamotrygon motoro* ryPP4a
44. *Drosophila melanogaster* pp4 (Y14213)
45. *Potamotrygon motoro* ryPP4b
46. *Paramecium tetraurelia* ppx1 (U31445)
47. *Saccharomyces cerevisiae* PPH3 (X82086)

PPG

48. *Saccharomyces cerevisiae* PPG (M94269)
49. *Schizosaccharomyces pombe* SPAC22H10.04 (Z69730)
50. *Trypanosoma brucei* TPP2 (M74168)

PP2A

51. *Homo sapiens* protein phosphatase 2A α (M36951)
52. *Potamotrygon motoro* ryPP2Aa

53. *Eptatretus burgeri* hgPP2A
54. *Potamotrygon motoro* ryPP2Ab
55. *Homo sapiens* protein phosphatase 2A β (X12656)
56. *Branchiostoma belcheri* amPP2A
57. *Drosophila melanogaster* PP2 (X78577)
58. *Ephydatia fluviatilis* sPP2A
59. *Neurospora crassa* pph-1 (X83593)
60. *Dictyostelium discoideum* PP2A (AF138278)
61. *Acetabularia cliftonii* PP2A (Z26654)
62. *Arabidopsis thaliana* PP2A-4 (U08047)
63. *Paramecium tetraurelia* pp2r-1 (AF030978)
64. *Schizosaccharomyces pombe* ppa2+ (M58519)
65. *Arabidopsis thaliana* PP2A (M96732)
66. *Schizosaccharomyces pombe* ppa1+ (M58518)
67. *Saccharomyces cerevisiae* PPH22 (X83276)

68. *Paramecium tetraurelia* PPN1 (U27497)

69. *Plasmodium falciparum* PP-beta (U89025)

70. *Arabidopsis thaliana* F21B7.27 (AC002560)
71. *Arabidopsis thaliana* T22O13.2 (AC007290)
72. *Arabidopsis thaliana* T4I9.4 (AF069442)
73. *Plasmodium falciparum* Pfpp-alpha (U88869)

- PPZ
86. *Saccharomyces cerevisiae* SAL6 (U00795)
87. *Saccharomyces cerevisiae* PPZ1 (M86242)
88. *Saccharomyces cerevisiae* ppz2 (L10241)
89. *Schizosaccharomyces pombe* phz1 (U73689)
90. *Neurospora crassa* pzl-1 (AF071751)

91. *Trypanosoma brucei* PP15.9 (X52746)

- PPY
92. *Drosophila melanogaster* protein phosphatase Y (Y07510)

- PP1
74. *Drosophila melanogaster* PPN 58A (Y17355)
75. *Arabidopsis thaliana* TOPP8 (U80922)
76. *Arabidopsis thaliana* TOPP2 (M93409)
77. *Arabidopsis thaliana* TOPP1 (M93408)
78. *Arabidopsis thaliana* TOPP4 (M93411)
79. *Arabidopsis thaliana* TOPP3 (M93410)
80. *Arabidopsis thaliana* AtPP1bg (Z46253)
81. *Acetabularia cliftonii* PP1 (Z28627)
82. *Acetabularia cliftonii* PP1 (Z28632)
83. *Chlamydomonas reinhardtii* PP1 (AF156101)
84. *Gonyaulax polyedra* PP1 (U52691)
85. *Schizosaccharomyces pombe* sds21+ (M27069)
93. *Arabidopsis thaliana* TOPP7 (U80920)
94. *Paramecium tetraurelia* PP1 (L26481)
95. *Dictyostelium discoideum* pppB (AF020537)
96. *Schizosaccharomyces pombe* dis2+ (M27068)
97. *Saccharomyces cerevisiae* DIS2S1 (M27070)
98. *Neurospora crassa* ppp-1 (AF124149)
99. *Ephydatia fluviatilis* sPP1c
100. *Drosophila melanogaster* PP1-beta 9C (X56439)
101. *Eptatretus burgeri* hgPP1b
102. *Homo sapiens* PPP1CB (X80910)
103. *Potamotrygon motoro* ryPP1b
104. *Ephydatia fluviatilis* sPP1a
105. *Ephydatia fluviatilis* sPP1b
106. *Drosophila melanogaster* PP1-alpha 96A (X56438)
107. *Branchiostoma belcheri* amPP1
108. *Homo sapiens* PPPICC (X74008)
109. *Potamotrygon motoro* ryPP1a
110. *Eptatretus burgeri* hgPP1a

付録4 本論文で解析に用いたギアルディアのPP遺伝子は、ギアルディアのDNA配列データベースから複数の配列を連結し、修正を加えたものである（材料と方法の章参照）。各遺伝子を構成するDNA配列の、データベースにおける番号を記した。配列の並びはN末側のものから順に示した。

Giardia gPP01
EI0184SA.SCF
NJ1942SA
EJ1226RA.SCF
AJ1785RA
EJ1226SA.SCF

Giardia gPP08
NG0270SA.SCF
MJ2587SA
NJ2656SA
MJ2587RA

Giardia gPP02
MI1905SA.SCF
EI1006RA.SCF
EI0124RA.SCF
MI1905RA.SCF
EJ3014SA

Giardia gPP09
KI1036RA.SCF
KJ2569RA
KJ2569RB
EJ4131RA
EJ2807RA
EJ4299SA
MD0836SA.SCF
EJ3690SA
EJ2807SA
EJ3690RA
KJ2569SA
EJ4299RA
KJ2569SB
EJ4131SA
NJ1041RA

Giardia gPP03
KJ2312RA
NJ2993SA
EJ3528RA
EJ2734RA
NJ2775SA
EJ3113RA
MD0125SA.SCF
EJ3113SA
AJ1442RA
NJ2993RA
KI0958SA.SCF
EJ3528SA
NJ2775RA
KI0676SA.SCF

Giardia gPP04
NI0945SA.SCF
NF0841SA.SCF
EJ3227RA

Giardia gPP05
MJ2453RA
NJ2361RA
EJ2311SA
HF2267RA.SCF
EI1743SA
MD0608SA.SCF
NJ2361SA
MJ2453SA
EJ2311RA
MD0608RA.SCF
HF2267SA.SCF
MD0715RA.SCF

Giardia gPP06
KJ2170SA
MJ2542RA
HF2660SA
NF0442RA.SCF
AI0580RA
MD0551RA.SCF
HF2660RA

Giardia gPP07
HF2701RA
HI2582RA
NJ2278RA
HF2701SA
MI1956RA.SCF
MI2005RA.SCF

付録5 系統樹推定に用いたアライメントを示す。系統樹推定に用いた領域のみを示し、それ以外の座位については、括弧内にその領域に存在するアミノ酸総数を示した。アミノ酸が保存している座位には*印を付した。

5-1 PTP遺伝子族のアライメント 番号は図7の分子系統樹上の番号に対応する

1.	(366)	NIGTIVMLGSLFEA(1)	REMCTAYWP(9)	YGDYCVKQISEE(2)	D(1)	SRFILRKFEI(9)	KKVHHYQYPNWSDCNSP(1)
2.	(505)	EVKVI13LNSDDEL(1)	LRKWDIYWN(17)	ICNINGCVLRVF(14)	D(1)	DLPHNGDLTS(8)	FIVYQLQYKNWLDSCGV(0)
3.	(623)	GVPLVL3LTNDFEN(1)	IEKCYRYWQ(3)	YNGIHVKLLLEKK(3)	M(1)	STTSMRKNTM(72)	YELLQIQYKNWPDGLTL(1)
4.	(115)	KSGIIVMLTKLRER(1)	RLKCDIYWP(8)	IGDLSVILVKVY(3)	S(1)	NEVQVREFEL(6)	KKILHFYNGWPDFGAP(1)
5.	(123)	DNIVIVMVTPLVEY(1)	REKCYQYWP(26)	PSDLKIEFVNVH(3)	D(1)	YTVTDIKLTP(8)	KTVHHFYFDLWKDMNKP(1)
6.	(983)	HLEIIVILTNLEEY(1)	KAKCAKYWP(8)	FGDILVKFAQER(0)	K(1)	GDYIERTLNV(13)	RQITQYHYLTWKDFMAP(1)
7.	(2)	NVPTIVMLTNIMEE(1)	KAKCSQYWP(7)	YGGLLVKVVETI(0)	I(1)	PDHTIRSLTV(7)	RTVLHFHYTNWHDYGT(1)
8.	(102)	KSRGVVMLNRVMEK(1)	SLKCAQYWP(10)	DTNLKLTLISED(0)	I(1)	SYTTRQLEL(8)	REILHFHYTTWPDFGVP(1)
9.	(102)	KSRGVVMLNRVMEK(1)	SIKCAQYWP(10)	DTNLKLTLISED(0)	I(1)	SYTTRQLEL(8)	REILHFHYTTWPDFGVP(1)
10.	(104)	KTKAVVMLNRIVEK(1)	SVKCAQYWP(9)	ETGFSVKLLSED(0)	V(1)	SYTVHLLQL(8)	RTISHFHYTTWPDFGVP(1)
11.	(2)	KSAGIVMLNKCTEK(1)	QAKCQYWP(8)	YNEFVVENLGE(0)	D(1)	GVYRLNRIKL(8)	RYISHFHYTHWPDFGVP(1)
12.	(122)	KSRVLMNLKLMK(1)	QIKCHLYWP(12)	HVKLTVELVRLE(0)	T(1)	QNFVRRWFKL(8)	REVMQFHYTTWPDFGIP(1)
13.	(2300)	KSTVIAMVTQVEG(1)	KIKCQYWP(9)	SNRLRALVRMQ(0)	Q(1)	KGFVVRAMTL(8)	RHISHLNFTAWPDHDT(1)
14.	(224)	NSNVIAMITREMEG(1)	IKCYHYWP(8)	LKHFRVFLENYQ(0)	I(1)	QYFIIRMFQV(8)	HSVKQLQFTKWPDHGTP(1)
15.	(735)	KLSLIVMLTTTLTER(1)	RTKCHQYWP(7)	HGGFHIQCQSED(0)	C(1)	IAYVSREMLV(8)	HTVTHLQYVWAPDHGIP(1)
16.	(744)	GSSMVVMLTTQVER(1)	RVKCHQYWP(7)	YGCYQVTCHEE(0)	G(1)	TAYIFRKMTL(8)	RPLTQIQYIAWPDHGVP(1)
17.	(1000)	GCVVIAMVTAEEEG(1)	RTKSHRYWP(10)	YGFKVTTLKFT(0)	D(1)	VCYATTGLKV(8)	RTVWHLQYTDWPDHGCP(1)
18.	(988)	GIATIAMVTAEEEG(1)	REKSFYWP(10)	YGRFKITTRFRT(0)	D(1)	GCYATTGLKM(8)	RTVWHLQYTDWPEHGCP(1)
19.	(2)	NVSIILMLTSEQEE(1)	QPKCHRYFP(10)	FEQYRITLKFVA(0)	Q(0)	NSVTTRCFSL(8)	REITHLQYSEWPDHGIP(1)
20.	(634)	KVSVIIVMLVSEAE(1)	KQKVARYFP(8)	HGALSALSLSSVR(0)	S(1)	ETHVERVLSL(8)	RSVLHLHFPTWPELGLP(1)
21.	(491)	KVLIIVMTTRVIER(1)	KIKCQYWP(8)	TGHFIIRNIHID(0)	L(1)	QDFKLTHLEV(8)	QSVAHYQYMSWPDFGVP(1)
22.	(394)	KVLIIVMTTRVIER(1)	RIKCGQYWP(8)	TGHFIIRNIHID(0)	L(1)	QDFKLTHFEV(8)	RSVAHYQYMSWPDFGVP(1)
23.	(393)	KVLIVIMTTREFE(1)	RRKCGQYWP(8)	FGFLTIVTNLGE(0)	N(1)	NHYKKTLEI(8)	RQVTHFQFLSWPDYGV(1)
24.	(2)	TVLIVIMTTKLVEA(1)	KVKCAEYWS(6)	HGGFIVQVTKVE(0)	K(1)	EGYEQRTMRV(6)	RTLAHFQFLAWPDYGP(1)
25.	(120)	SVLIIGMACMEYEM(1)	KKCCERYWA(8)	FGPFSVSCAEK(0)	R(1)	SDYIIRTLKV(6)	RTIYQFHYKSWPDHVP(1)
26.	(124)	NVVIIVMACREFEM(1)	RKKCERYWP(8)	FAPFKISCDEEQ(0)	A(1)	TDYFIRTLLE(6)	RRLYQFHYVNWPDHVP(1)
27.	(122)	GVKIVMLACREIEN(1)	RKKCERYWA(7)	TGLFCITLIEK(1)	L(1)	EDIMLRTLKV(6)	RSVYQLQYMSWPDGVP(1)
28.	(2)	RVKCVVMTTNIIEK(1)	SKKCEKYWP(7)	FGAITVSGIQK(0)	Q(1)	EGFILSKLTL(8)	HQVYHFHYTLWPDFGSP(1)
29.	(801)	GCTVIVMLTPLVED(1)	VKQCDRYWP(6)	YHVVENVLVSEH(1)	W(1)	EDFLVRSFYL(8)	RTLTQFHFSLWPDFGTP(1)
30.	(837)	GCVVIVMLTPLAEN(1)	VRQCXYHYWP(6)	YHIEVENVLVSEH(1)	W(1)	EDFLVRSFYL(8)	RTVTQFHFSLWYDRGVP(1)
31.	(348)	NSRVIVMTTKEVER(1)	KSKCVKYWP(7)	YGVMRVRNVKES(0)	A(1)	HDYTLRELKL(9)	RTVWQYHFRTWPDHGVP(1)
32.	(348)	NSRVIVMTTKEVER(1)	KSKCVKYWP(7)	YGVMRVRNVKET(0)	P(1)	HDYILRELKL(9)	RAVWQYHFRTWPDHGVP(1)
33.	(470)	NTRVIVMTTKEVER(1)	NEKCARLYWP(7)	FGHARILQVSEN(0)	S(1)	SDYTLREFLV(7)	RRIFHYHFQVWPDHGVP(1)
34.	(2)	NSLIIVMTTNEVER(1)	RNKCTRYWP(7)	YGKVTVCCVKET(0)	T(1)	PHYVLREFV(14)	RTIFQYHFKAHPDHGVP(1)
35.	(303)	NSRVIVMTTREVEK(1)	RNKCVPYWP(7)	YGPYSVTNCGEH(0)	D(1)	TEYKLRTLQV(9)	REIWHYQYLSWPDHGVP(1)
36.	(362)	HTPIIVMITNIEEM(0)	NEKCTEYWP(5)	YDGEIVTVQKVI(0)	H(1)	EDYRLRLISL(6)	RGLKHYWFTSWPDQTP(1)
37.	(483)	DSPVIVMTTKLEK(0)	NEKCVLYWP(5)	YDGFLVTQKSVQ(0)	E(1)	DNYTIRNLVL(6)	QHVKKHYFTSWPDHKT(1)
38.	(186)	EVSLIVMLTQLREG(0)	KEKCVHYWP(5)	YGPFIQIRIQDMK(0)	E(1)	PEYTVRQLTI(6)	RSVKHILFSWPDHQT(1)
39.	(2)	NVPSIVMLTKLVEN(1)	KEKCSRYWP(9)	YGQYSIRNFEQ(0)	S(1)	ASYTISYLT(10)	KDVVHLWYTAWPDFGVP(1)
40.	(335)	NVEVIVMITNLLEK(1)	RRKCDQYWP(6)	YGNFLVTQKSVH(0)	V(1)	AYYTVRNFTL(16)	RVVTQYHYTQWPDGMVP(1)
41.	(1817)	NVEVIVMITNLLEK(1)	RRKCDQYWP(6)	YGNFLVTQKSVQ(0)	V(1)	AYYTVRNFTL(16)	RVVTQYHYTQWPDGMVP(1)
42.	(942)	NTGIIVMITNLVEK(1)	RRKCDQYWP(6)	YGNIIIVTLKSTK(0)	I(1)	ACYTVRRFSI(19)	RVVIQYHYTQWPDGMVP(1)
43.	(919)	HTGIIVMITNLVEK(1)	RRKCDQYWP(6)	YGNIIIVTLKSTN(0)	I(1)	ACYTVRPLHG(19)	RTVIQYHYTQWPDGMVP(1)
44.	(569)	RVAIIVMITNLVER(1)	RRKCDMYWP(6)	YGVIVQVKLIEE(0)	V(1)	STYTVRTLQI(14)	KLIVQYHYTNWPDHGT(1)
45.	(741)	KATVIVMVTTRCEE(1)	RNKCAEYWP(8)	FGDILTVRISEEQ(0)	R(1)	PDYIIQKLN(9)	REVTHIYFTSWPDHGVP(1)
46.	(641)	KATVIVMVTTRCEE(1)	RPKCAQYWP(8)	FGDILTVRISEEQ(0)	W(1)	PDYVIRKFLI(9)	REVTHIQFIRWPDHGVP(1)
47.	(225)	KSATIVMLTNLKER(1)	EKCKHQYWP(6)	YGNIRVVCVEDCV(0)	V(1)	VDYTVRKFCI(11)	RLVSQHFHTSWPDFGVP(1)
48.	(331)	NTATIVMVTNLKER(1)	EKCKAQYWP(6)	YGNIRVSVEDVT(0)	V(1)	VDYTVRKFCI(12)	RLITQHFHTSWPDFGVP(1)
49.	(351)	NTATIVMVTNLKER(1)	EKCKAQYWP(6)	YGNIRVSVEDMT(0)	V(1)	VDYTVRKFCI(12)	RLITQHFHTSWPDFGVP(1)
50.	(994)	NSISVVMVTNLVEV(1)	RVKCVRYWP(5)	YGDIVKVSIAIETE(0)	P(1)	AEYIIRTFTV(8)	REIRQHFHTSWPDNGVP(1)
51.	(998)	NSASIVMVTNLVEV(1)	RVKCVRYWP(5)	YGDIVKVTLIETE(0)	P(1)	AEYVIRTFV(8)	RELRLHFHTSWPDHGVP(1)
52.	(989)	NTASIVMVTNLVEV(1)	RVKCKYWP(5)	YKDIKVTLIETE(0)	L(1)	AEYVIRTFV(8)	REIRQHFHTSWPDHGVP(1)
53.	(976)	QSASIVMVTNLVEV(1)	RVKCYKYP(5)	YGDIVKVTCEME(0)	P(1)	AEYVIRTFV(8)	REIRQHFHTSWPDHGVP(1)
54.	(969)	HCSSIVMITKLVEV(1)	RVKCSRYWP(5)	YGDIVKIMLVKTE(0)	T(1)	AEYVIRTFV(8)	HEVRQHFHTAWPEHGVP(1)
55.	(2)	DCPTIVMVTNLQEE(1)	KVKCHQYWP(6)	YGHILVTLKEVE(0)	N(1)	AEYSIRTFV(13)	REVRFHFHTWPDHGVP(1)
56.	(1563)	KTATIVMMTRLEER(1)	RIKCDQYWP(6)	YQIFVTTITETQ(0)	E(1)	ATYSIRTFQL(8)	REIKQLQFTAWPDHGVP(1)
57.	(964)	RVTIVMVTMKLEER(1)	RIKCDQYWP(6)	YGPDMHVTLVDDV(0)	E(1)	AIYTLRTFIV(9)	IEVKQFHTWPDHGVP(1)
58.	(1446)	RSATVVMVTMKLEER(1)	RVKCDQYWP(6)	HGLVQVTLTLDIV(0)	E(1)	ATYCVRTFAL(8)	REVRFQFHTWPDHGVP(1)
59.	(1431)	RTATVVMVTMKLEER(1)	RVKCDQYWP(6)	CGLIQVTLTLDIV(0)	E(1)	ATYTVRTFAL(8)	RELRFQFHTWPDHGVP(1)
60.	(1034)	RSATVVMVTMKLEER(1)	RIKCDQYWP(6)	YGMIVQVTLTLDI(0)	E(1)	ATFCVRTFSL(8)	REVRFQFHTWPDHGVP(1)
61.	(1482)	RSATVVMVTMKLEER(1)	RIKCDQYWP(6)	YGFIVQVTLTLDI(0)	E(1)	ATFCVRTFSL(8)	REVRFQFHTWPDHGVP(1)
62.	(1578)	QSPVIVMVTMKLEER(1)	RVKCDQYWP(6)	YGSILQVNLVDVI(0)	E(1)	ATYTMRTFQM(9)	REVRFQFHTWPDYGV(1)
63.	(2)	GSATIVMLTNLKEE(1)	RIKCNQYWP(6)	YGSIRIHTQEE(0)	T(1)	ADYTLRIFSL(8)	RIVKQFHTWPDHVP(1)
64.	(2)	QLQTIIVMLTKCSEG(1)	KIKCEMYWS(6)	TDTLTVTTTSEE(0)	S(1)	ADFQIRKFSI(10)	LQVTHFHYTAWPDHGVP(1)
65.	(1199)	NIYSIVMLTKCVEQ(1)	RTKCEQYWP(6)	YGDIVVTMSEV(0)	V(1)	PEWTIRDFNV(9)	HTVRQHFHTSWPDHGVP(1)
66.	(1130)	NVYAIIMLTCKVEQ(1)	RTKCEEYWP(6)	YGDITVAMTSEI(0)	V(1)	PEWTIRDFTV(9)	HPLRQHFHTSWPDHGVP(1)
67.	(1793)	NVHNIVMVTQCVEK(1)	RVKCDHYWP(7)	YGDILQLMLSES(0)	V(1)	PEWTIREFKI(9)	RLIRHFHYTVWPDHGVP(1)
68.	(1362)	NSRAIVMLTRCFEK(1)	REKCDQYWP(7)	YGDIVKQILNDS(0)	H(1)	ADWVMEFML(6)	RILRHFHTWPDHGVP(1)
69.	(1419)	NSRAIVMLTRCFEK(1)	REKCDQYWP(7)	YGDIVKQILNDS(0)	H(1)	ADWVMEFML(6)	RILRHFHTWPDHGVP(1)
70.	(1028)	KSQIVMLTQCNEK(1)	RVKCDHYWP(7)	YGDIVTEMISEE(0)	E(1)	DDWACRHFRI(6)	QDVMHFNYTAWPDHGVP(3)
71.	(1079)	KSQIVMLTQCNEK(1)	RVKCDHYWP(7)	YGDIVTEMISEE(0)	E(1)	DDWACRHFRI(6)	QDVMHFNYTAWPDHGVP(3)
72.	(2096)	RTKTLVMLTQCFEK(1)	RIRCHQYWP(8)	FGDIVITKLMED(0)	I(1)	IDWTIRDLKI(6)	MTVRQCNTGWPEHGVP(1)
73.	(1240)	QVHVIMMLTQGMEN(1)	RVLCHEHYWP(7)	HGHIHTHLLAE(0)	S(1)	DEWTRREFQL(8)	RRVKQLQFTWPDHGVP(1)
74.	(912)	QSHTLVMLTNLCEA(1)	RVKCEHYWP(7)	HGHLRVTLVGEE(0)	V(1)	ENWTVRELL(9)	LSVRCFHYQAWPDHGVP(1)

5-2 カワカイメンのPTPR3サブタイプに属する遺伝子のアライメント。-はギャップを表す。一番上がsPTPR3遺伝子である。

```
( 0) EGQLQTIIVMLTKCSEGGKIK-CEMYWSDSMGGVYE-TDTLTVTTTSSNESYADFQIRKFSIKCKVA
( 0) QEKTNTIIVMLTQCNEGGKRIK-CEQYWPEKVGETIEPFPGLKVTLNDFMTFADYHTRKMSILNN-G
( 0) EGRQLQTIIVMLTKCFEGGKIK-CEMYWCDGLGGVYD-ADDLTVITTSIQSFADFGIRKFSVCSRVA
( 0) EKHIETIIVMLTRCNEAGKSI-CKQYWTDEEGGVYD-TDTLSMATTSSVVTLADYDIRTFFEVKSKME
( 0) EKRLPTIVVMLTELMEGGRSK-CERYWPENINDNWDVGCNLRVCLIEQRPFAEYTVKILVVTDTS
( 0) EKQIETIIVMLTRCNEAGKSM-CKQYWTDEEGGVYD-TDTLSMATTFFAMTLADYDIRTFFEVKSKME
( 0) ENQVQTIIVMLTKCIESGKKK-CEPYFSDVPGGTFE-TEAMVVTTTSIVPYADFEIKKLTIVKSKTE
( 0) EGRLQTIIVMLTKCFEGGKIK-CEMYWSNRVGGVYE-TDDLIVTTTSIQSFADFGIRKFSIISKVA
( 0) EGRQLQTIIVMLTKCFEGGKIK-CEMYWCDGMGGVYE-TDDLTVTTTSIQSFADFGIRKFSVCSKVA
( 0) EKQIETIIVMLTRCNEAGKSM-CKQYWTDEEGGVYD-TDTLSMTTTSVVTLADYDIRTFFEVKSKME
( 0) EGRQLQTIIVMLAKCTEGGKHDCCEMYWSDSMGSGVYE-IDTLTVTTTISILLFADFDIRKFSIKNKMA
( 0) ESRLQTIIVMLTKCFEGGKIK-CEMYWCDGMGGVYE-TDDLTVTTTSIQSFADFGVRKFSVSSKVA
( 0) DKRYPTIIVMLTELIEGGKVK-CERYWPEHVNESWDVGHNLRVTLLEQKPFVEYRLKILTVDISS
( 0) EGQLQTIIVMLTKCSEGGKDK-CEMYWSDSMGEGVYE-TDTLTVTTTISIVPFADFDIRKFSVSKSLA
( 0) ESRLQTIIVMLTKCFEGGKIK-CEMYWCDVMGGVYE-ADDLTVTTTSIQSFADFGVRKFSVSSRVA
( 0) EKHIETIIVMLTRCNEAGKSM-CKQYWTDEEGGVYD-TDTLSMTTTSAMTLADYDIRTLEVKSKME
( 0) EYKLTIVVMLTKCYEDEKAK-CECYWPLEVGQSVEYGAHIVTLQSVVMCTEYDIRTLTVTQVDN
( 0) EGQLKTIIVMLTKCLEGGKIK-CEMYWSDRVGGVYE-TDDLTVTTTSIQSFADFGIRKFSIISKVA
      *  *  *  *  *  *  *  *
```

```
PDADCLQVTHFHYTAWPDHGVPPQTATSILSFVRRVQNAHDKGKGVPLL ( 71)
SEQ--LTISHFQFLGWPDHGVPHYSTSLLEFIKVRNEHKK-DGSPLL (371)
PDQDKLEVTFQHYTSWPDHGVPPQFATSFLKFVRRRIQKAHDKSKGVPLL ( 0)
EESDVLKVTHYQYTSWPDDEVPPQFATSFLNFVRRVQKTNDKSKGVPLL ( 0)
SDAPSLRVTHYHFTVWPDHGVPPADKTCMIQFIKVRNTHPY-KGPPLV (364)
EDSDVLKVTHYQYTSWPDDEVPPQFATSFLNFVRRVQKTHDKSKGVPLL ( 0)
PDSEPLKITHYHFTSWPDHGVPPQFATSISFVRKVQKTHDKSKGVPLL ( 0)
PDQDKLEVTFQHYTSWPDHGVPPQFATSFLKFVRRRIQSSHDKSKGVPLL ( 0)
PDQDKLEVTFQHYTSWPDHGVPPQFATSFLKFVRRRIQKAHDKSKGVPLL ( 0)
GDSDDLKVTHYQYTSWPDNEVPKFATSFLNFVRRVHKAHDKSKGVPLL ( 0)
PDTDPLQVTHFHTAWPDHGVPPQFATSILNFLRRVQNSHNKGKGVPLL ( 0)
PDQDKLVVTFQHYTSWPDHGVPPQFATPFLKFVRRRIQKAHDKSKGVPLL ( 0)
PTASPLRINHYHFTVWPDHGAPADKTCMIQFIKVRNIHPYSDANPLVI ( 0)
PDAGYLQVTFQHYTAWPDHGVPPQFATSLLSFVRKVQKTQDKSKGVPLL ( 0)
PDQDKLEVTFQHYTSWPDHGVPPQFATSFFKFVRRRIQKAHDKSKGVPLL ( 0)
EDSDVLKVTHYQYSSWPDNEVPKFVTSFLNFVRRVHKAHNKSKGIPLL ( 0)
TDQLPLTVTQYQFLGWTDHGVPEYATSLTTFIRHVQR--RPVGVPM ( 0)
PDQDKLEVTFQHYTSWPDHGVPPQFATFILKFVRRIRSAHEKSKGVPLL ( 0)
      *  *  *  *  *  *
```


5-3 PTPR4サブタイプのアライメント。番号は図13の分子系統樹上の番号に対応する。

```

1. ( 0)EQNTATIVMVTNLKERKECKCAQYWP (0)DQGCWTY (0)GNIRVSVEDMMVLVDYTIKRFCEI (12)RLVTQHFHTSWPDPFGVPFT
2. (329)EQNTATIVMVTNLKERKECKCAQYWP (0)DQGCWTY (0)GNIRVSVEDVTVLVDYTVRKFCI (12)RLITQHFHTSWPDPFGVPFT
3. (335)EQNTATIVMVTNLKERKECKCAQYWP (0)DQGCWTY (0)GNIRVSVEDVTVLVDYTVRKFCI (12)RLVTQHFHTSWPDPFGVPFT
4. (349)EQNTATIVMVTNLKERKECKCAQYWP (0)DQGCWTY (0)GNIRVSVEDMTVLVDYTVRKFCI (12)RLITQHFHTSWPDPFGVPFT
5. ( 0)EQNTATIVMVTNLKERKECKCAQYWP (0)DQGCWTY (0)GSIRVAVEDVTVLVDYTIKRFCEI (13)RLVTQHFHTSWPDPFGVPFT
6. (223)EQKSATIVMLTNLKERKECKCHQYWP (0)DQGCWTY (0)GNIRVCVEDCVVLVDYTIKRFCEI (11)RLVSQHLFTSWPDPFGVPFT
7. ( 0)EQKSATIVMLTNVKEKKECKCYKYP (0)DQGLWTY (0)GAIRVSVEDVTVLVDYTIKRFCEI (10)RLVTQLHFTSWPDPFGVPFA
8. ( 0)EQNTATIVMVTNLVKEKNKVKCTQYWP (0)DTGSKQY (0)GDITVRSEETSTLVDYIIRTLTL ( 6)RTLHHEFTTWPDPFGVPKS
9. ( 0)ETGSTAIVMVTNLEENGKKKCSQYWP (0)TSGEKIY (1)ADLVVRLEETVKMVDHVTRTFLL ( 8)RKVTHEHFIGWPDGFLPKS
10. ( 0)ETGSTVIVMVTNLEEKGKTCTQYWP (0)ESGEEVY (0)QGISVTLAETIPMYYVTVRVLLV ( 7)RKITQFQFLGWPDGAPRN
11. (739)EQKATIVIMVTRCEEGRNRKCAEYWP (2)EEGTRAF (0)GDVVVKINQHKRCPDYIIQKLNI ( 9)REVTHIQFTSWPDHGVPE
12. (674)EQKATIVIMVTRCEEGRNRKCAQYGP (2)ENGSAQY (0)GDITVKINESKICPDYIIQKLHI ( 9)RDVTHIQFTSWPDHGVPE
13. (639)EQKATIVIMVTRCEEGRKPKCAQYWP (2)DSPSKTF (0)GDLTVRISEEQWCPDYIVIRKLF ( 9)REVTHIQFIRWPDHGVPE
*      *** *      *      *      *      *

```

```

1. FIGMLKFLKKVKTCPNPSFAGAIIVHCSAGVGRGTGTFIVIDAMLDMHHAERKVDVFGFVSRIARQRCQMVTDMQYVFYQALLEHYLYGD
2. FIGMLKFLKKVKACNPQYAGAIIVHCSAGVGRGTGTFVVIDAMLDMHHTERKVDVYGFVSRIARQRCQMVTDMQYVFYQALLEHYLYGD
3. FIGMLKFLKKVKTCPNPSFAGAIIVHCSAGVGRGTGTFIVIDAMLDMHHAERKVDVYGFVSRIARQRCQMVTDMQYVFYQALLEHYLYGD
4. FIGMLKFLKKVKACNPQYAGAIIVHCSAGVGRGTGTFVVIDAMLDMHHTERKVDVYGFVSRIARQRCQMVTDMQYVFYQALLEHYLYGD
5. FIGMLKFLKKVKTCPNPSYSGPIVHCSAGVGRGTGTFIVIDAMLDMIAEHKVDVFGFVSRIARQRCQMVTDMQYVFYQALLEHYLYGD
6. FIGMLKFLKKVKTCPNPSYSGPIVHCSAGVGRGTGTFIVIDAMLDMIAEHKVDVFGFVSRIARQRCQMVTDMQYVFYQALLEHYLYGD
7. FIGMLKFLKKVKALNPSYSGPIVHCSAGVGRGTGTFIVIDAMLDMIAEHKVDVFGFVSRIARQRCQMVTDMQYVFYQALLEHYLYGD
8. PLGMMKFVRRVKAANPADSGPIVHCSAGVGRGTGTFIVIDAMFEMIAAEQKVDVFGFVSRIARQRCQMVTDMQYVFYQALLEHYLYGD
9. PMGLLKHFRTVMSALTARDRPIVHCSAGVGRGTGTFITVDAMLDMIRAEKVDVFGFVEEMRQNRSMVQTEGQYVFYQALLEHYLYGD
10. PAELLKFHQMVTSTTSRDRPIVHCSAGVGRGTGTFITVDAMLAMMAEKGVDVFGFVSRIARQRCQMVTDMQYVFYQALLEHYLYGD
11. PHLLKLRRRVNALSNNFSGPIVHCSAGVGRGTGTYIGIDAMLEGLEAENKVDVYGYVVKLRQRCLMVQVEAQYILIHQALVEYNQFGE
12. PHLLKLRRRVNALSNNFSGPIVHCSAGVGRGTGTYIGIDAMLEGLEAENKVDVYGYVVKLRQRCLMVQVEAQYILIHQALVEYNQFGE
13. PHLLKLRRRVNALSNNFSGPIVHCSAGVGRGTGTYIGIDAMMQGLEAENKVDVYGYVVKLRQRCLMVQVEAQYILIHQALVEYNQFGE
*      *      *      *      *      *      *      *      *      *      *      *      *

```

```

1. TELEVTSLESHLQKLYNRVS (1)SGCNGLEEEFKKLTSIKIQNDK (1)RTGNLPANMKKNRVLQIIPYEFNRVII ( 8)TDYVNASFIDG
2. TELEVTSLETHLQKIYNKIP (1)TSNNGLEEEFKKLTSIKIQNDK (1)RTGNLPANMKKNRVLQIIPYEFNRVII ( 8)TDYVNASFIDG
3. TELEVTSLEIHLQKIYNKVP (1)TSNNGLEEEFKKLTSIKIQNDK (1)RTGNLPANMKKNRVLQIIPYEFNRVII ( 8)TDYVNASFIDG
4. TELEVTSLETHLQKLYSKFP (1)TNSTGLEEEFKKLTSIKIQNDK (1)RTGNLPANMKKNRVLQIIPYEFNRVII ( 8)TDYVNASFIDG
5. TELEVTSLESHLQKLYSKQP (0)SGINGLEEEFKKLTVKIQNDK (1)RTGNLPANMKKNRVLQIIPYDFNRVII ( 8)TDYVNASFIDG
6. TELDVSSLEKHLQTMHGTFTT (1)FDKIGLEEEFKKLTVNRIMKEN (1)RTGNLPANMKKARVLIIPYDFNRVIL ( 8)TDYVNASFIDG
7. TELDVSSLEKHLQKLRNKTG (1)FDITGLEEEFKKLTVNRIMKEN (1)RTGNLPANMKKARVLIIPYDFNRVIL ( 8)TDYVNASFIDG
8. TEIEVTLNRRHLQQLAAKLP (1)SQDTGMEAEFKKLTVNRIMKEN (1)RSGNMADNIKKNRVLQILPYDTSRVYI ( 8)SDYINAAFDVG
9. TETEVANIHYRMQKKAEDP (1)TGKTGIEFEGKLTRIPIDKAN (1)RSGNLPENLSKNRVLQVLPYDTSRVFL ( 8)ADYINASFIDG
10. TEVEVTSIHRHLQNLRRREDQ (1)SQKTGLEELEYEKLTRVPVDKAN (1)RNGNLPENLSKNRVLQVLPYDTSRVFL ( 8)SDYINASFIDG
11. TEVNLSLHLPYLNHMKKRD (1)SEPSPLEAEFQRLPSYRSWRTQ (0)HIGNQEENKSKNRNSNVIPIYDYNRVPL (27)SKYINASFIMS
12. TEVSLSELHLYNNLKRKDP (1)SEPSLLEAKFQRLPSYKGWRTQ (0)NTGNREENKSKNRNSANTIPYDFNRVPI (27)SKYINASFIMS
13. TEVSLSELPHKLINFKKNP (1)SEPSMLEGEFQRIPIYTDWRTQ (0)TTGRRGENQSKNRSLSVIADYNRVTI (28)TKYINASYIDG
**      *      *      *      *      *      *      *

```

```

1. YRQKDS (1)IASQGGLQHTIEDFWRMIWEWKSCSIVMLTELE (2)GQEKCAQYW (1)SEGTV (1)YGDITIEIKKEESES YTVRDLVL (
2. YRQKDS (1)IASQGGLLHTIEDFWRMIWEWKSCSIVMLTELE (2)GQEKCAQYW (1)SDGLV (1)YGDITVELKKEESES YTVRDLVL (
3. YRRRTP (0)TCQPRPVQHTIEDFWRMIWEWKSCSIVMLTELE (2)GQEKCAQYW (1)SDGSV (1)YGDINVELKKEESES YTVRDLVL (
4. YRQKDS (1)IASQGGLRHSLEDFWRMIWEWKSCSIVMLTELE (2)GQEKCAQYW (1)SEGTM (1)FGDITIELKKEESES YTVRDMFV (
5. YRQKDA (1)IAGQGGLSHTLEDFWRMVWEWNSCSIVMLTELE (2)GQEKCTQYW (1)NEGTL (1)SGDISIELKSETHHDS YTVRDFGV (
6. YRQKDY (1)IATQGGLAHTVEFWRMIWEWKSHITIVMLTEVQ (2)EQDKCYQYW (1)TEGSV (1)HGEITIEIKNDTSEASIRDFLV (
7. YRQKDY (1)VATQGGLAHTVEFWRMVWDCKSHITIVMLTELQ (2)EQEKCCQYW (1)SEGSV (1)YGTYSTELKKNITDAFSLRDLTV (
8. YREKDA (1)IATQGGLPHTVNDFWKMWWEWKSCSIVMLTELQ (2)GHEKCHKY (1)GEVEM (0)YGDICVEAKGDKTFQDYTVRTHFI (
9. YNMKDA (1)IATQGGLDRTEFWRMVWEWNSCSIVMLTEIK (2)SQNKCAMY (1)EDGSL (1)SGDFTVSLQDQTDYQDYTLRTHFL (
10. YREKDA (1)IATQGGLDRTEFWRMVWEWNSCSIVMLTELQ (2)QRSTCVRY (1)QEGTK (1)YGDITVRLDDTSEFGDFTHRTFSL (
11. YWKPEV (1)IAAQGLPKETIGDFWQMIQFQRKVKVIVMLTELK (2)DQEICAQYW (0)GEGKQ (1)YGDIEVDLKDTSSTYTLRVFEL (
12. YWGPKA (1)IATQGGLQETISDFWQMVFGQRKVKVIVMLTELK (2)DQELCAQYW (0)GEGRQ (1)YDDIEVQVTDVNCPSYTIRAFV (
13. YWHSET (1)IATQTPLEPETIADFWMVMVYQKARIAMLGK (0)DDKDCSQYW (0)EDDKK (1)YDDIEVVLSECNKQPEFTVIRIFEI (
*      *      *      *      *      *      *

```

5-3 続き

```

1. 8) RQIRQFHFHGWPEVGIPADGKGMINLIAAV(0) QK( 8) PITVHCSAGAGRTGTFCALSTVLERVKAEGILDVFQTVKSLRLQRPHM
2. 8) RQIRQFHFHGWPEVGIPSDGKMISIIAAV(0) QK( 8) PITVHCSAGAGRTGTFCALSTVLERVKAEGILDVFQTVKSLRLQRPHM
3. 8) RQIRQFHFHGWPEVGIPSDGKGMINIIAAV(0) QK( 8) PMHCHCSAGAGRTGTFCALSTVLERVKAEGILDVFQTVKSLRLQRPHM
4. 8) RQIRQFHFHGWPEVGIPDTGKGMINIIAAV(0) QK( 8) PITVHCSAGAGRTGTFCALSTVLERVKAEGILDVFQTVKSLRLQRPHM
5. 8) RQIRQFHFHGWPEVGIPDDGKGMIGLIAAV(0) QK( 8) PIIVHCSAGAGRTGTFCALCTILERVKTEGILDVFQTVKSLRTQRPHM
6. 13) RVVRQFHFHGWPEIGIPAEGKGMIDLIAAV(0) QK( 8) PITVHCSAGAGRTGTFCALCTILERVKTEGILDVFQTVKSLRTQRPHM
7. 8) RTVRQFHFHGWPEIGIPAEGKGMIDLIAAV(0) QK( 8) PIIVHCSAGAGRTGTFCALCTILERVKTEGILDVFQTVKSLRMQRPHM
8. 26) RTVQQFHFHGWPEIGIPANAAGMLDLIGQV(0) ER( 8) PITVHCSSGAGRTGAFITLSTVIERVKAEGICDVFQTVKSMRYQRPHM
9. 8) RTVQQFHFHGWPEVGIPDNASGMIDLIGQV(0) QK( 8) PITVHCSAGAGRTGAFCAISTVLERVKAEGVCDVFQVVKALRLQRPHM
10. 9) RTIHQFHFHGWPEQYGVDPNAAGIIDLIGQV(0) QK( 8) PITVHCSSGAGRTGTFCALSTVLERVKAEGVCDVFQVVKALRLQRPHM
11. 8) RTVYQYQYTNWSVEQLPAEPKELISMIQVV(1) QK(16) PLLIHCRDGSQQTGIFCALNLLESAETEEVVDIFQVVKALRKARLGM
12. 8) QKVYQYQYHKWNGLDVPEDPKDLVDMILSL(1) QK(16) PFVIHCCDGSQQTWCVLCLMTLLESAETEEVIDVFQVVKALRRSRLGV
13. 8) RQVYQYHFDWAESSELPEDPSNFTKMIRSI(1) EK(13) SLIVHCSGAKKTGVFYALWILLDNADTENVIDVLQTVKVLRKARPGL
      *      *      *      *      *      *      *      *      *      *      *      *

```

```

1. VQTLEQYEFQYKVV(14)
2. VQTLEQYEFQYKVV(14)
3. VQTLEQYEFQYKVV(14)
4. VQTLEQYEFQYRVV(14)
5. VQMLEQYEFQYKVV(14)
6. VQTLEQYEFQYKVV(14)
7. VQTLEQYDFQYKVV(14)
8. VQTVEQYQFYQAV(17)
9. VQTLQYQFYQAV(14)
10. VQTIDQYQFYQAV(14)
11. VSTFEQYQFLYDVI(81)
12. VSTFEQYQFLYDTI(79)
13. VSTFEQYQFLYDII(81)
      *      *      *

```


5-4 PTPN3サブタイプのアライメント。番号は図14の分子系統樹上の番号に対応する。

```
1. (100)EQKSRGVVMLNRVMEKGS LKCAQYWP ( 6)MIFE(1)TNLKLTLISEDIKSY YTVRQLELENLTQETREILHFHYTTWPDFGVP
2. (100)EQKSRGVVMLNRVMEKGS IKCAQYWP ( 6)MFFE(1)TNLKLTLISEDIKSY YTVRQLELENLTQETREILHFHYTTWPDFGVP
3. (102)QKTKAVVMLNRIVEKESVKCAQYWP ( 5)MLFK(1)TGFSVKLLSE DVKSYYTVHLLQLENINSGETRTISHFHYTTWPDFGVP
4. ( 0)EQKSRAIVMLNRVIEKGSVKCEKYWP ( 6)MSYE(1)TGFNLKLLCEDVKSYYTVRRLELQRTSTGETREIYHFHYTTWPDFGVP
5. ( 0)EQQSRAVVMLNRVLEKGLVKCEQYWP ( 6)MVFD(0)TGYRLTLASSIPKEYFTISR LMLECLSTAERREVLHFHYTSWPDFGVP
6. ( 0)EQRTKAVVMLNRVIEKGTIKCAQYWP (10)LELE(1)TGYKVTLL EEDVRPYFTIRTYLLQRLKTNESRKVFHFHYTRWPDFGVP
7. (120)EQKSRAVLMNLNKLMEKKQIKCHLYWP ( 8)LKLP(1)VKLTVELVRLE TYQNFVRRWFKLTDLETQQSREVMQFHYYTTWPDFGIP
      *          ***      *      *      *      *
      *          ***      *      *      *      *

1. ESPASFLNFLFKVRESGSLSP EHGPPVVHCSAGIGRSGTFCLADTCLLLMDKRKD(3)VDIKKV LLEMRRKFRMG LIQTADQLRFSYLAVI
2. ESPASFLNFLFKVRESGSLNPEYGPVVVHCSAGIGRSGTFCLVDTCLLLMDKRKD(3)VDVKQV LLEMRRKFRMG LIQTADQLRFSYLAVI
3. ESPASFLNFLFKVRESGSLNPDHGPAVIHCSAGIGRSGTFSLVDTCLVLMEKGDD(0)INIKQV LLNMRKYRMGLI QTPDQLRFSYMAII
4. ESPASFLNFLFKVRESGCLNSDCGPTIVHCSAGIGRSGTFSLVDTCLVLMEKRKD(3)VDIREV LLNMRKYRMGLI QTPDQLRFSYMAVI
5. DSPAAFLAFLMRVRRSGVLDPSFGPIVVHCSAGIGRSGTFCLVDSCLALHAARGE(0)PDVYQV LLDLRRFRMG LIQTPAQLRFSFLAII
6. ESPAAFLHFLLEVRDSGSLES DVGPPVVHCSAGIGRSGVFCVDTCLVLLDKTRD(3)LNIRRL LLDMRQYRMGLI QTPDQLRFSYLAVI
7. SSPNAFLKFLQQVRDSGCLSRDVGPAVVHCSAGIGRSGTFCLVDCCLVLIDKYGE(0)CNVSKV LCELRTYRMGLI QTDQLDFSQAI I
      **  **  **  **  **  *      **  *****  *  *  *  *  *
      *          *          *          *          *          *          *          *          *          *

1. EGAKFIMGDSSVQDQ(145)
2. EGAKFIMGDASVQEQ(144)
3. EGAKCIKGDSSIQKR(127)
4. EGAKYIMGDSSVQKR(132)
5. QGMHLLSADACQEG(193)
6. EGARTILGMTTAEDE(137)
7. EGIKKLHDPTFLDAE(239)
      *
```

5-5 PTPR5サブタイプのアライメント。番号は図15の分子系統樹上の番号に対応する。

```

1. (1815) EHNVEIVMITNLVEKGRRKCDQYWPADGSEEEYGNFLVTQKSVQVLAYYTVRNFRTLNTKIKKG ( 8) RVVTQYHYTQWPDGMV
2. ( 333) EHNVEIVMITNLLEKGRKCDQYWPAGSEEEYGNFLVTQKSVHVLAYYTVRNFRTLNTKIKKG ( 8) RVVTQYHYTQWPDGMV
3. ( 0) EHNVSIVMITNLVEKGRRKCDQYWPSENSEEEYGHILVSLKNTKVSSYYTVRQFTVRNTKFKKG ( 8) HVTQYHYTQWPDGMV
4. ( 940) EQNTGIIIVMITNLVEKGRRKCDQYWPTESENSEEEYGNIIIVTLKSTKIHCYTVRRFSIRNTKVKKG (11) RVVIQYHYTQWPDGMV
5. ( 917) AQHTGIIIVMITNLVEKGRRKCDQYWPTESENSEEEYGNIIIVTLKSTNIHCYTVRPLHGQEHKDEKG (11) RTVIQYHYTQWPDGMV
6. ( 0) EQNTGIIIVMITNLVEKGRRKCDQYWPSENSEEEYGNIIIVTLKSTKVRACCTVRHFTIRNTKVKKG (11) RMVIQYHYTQWPDGMV
7. ( 0) EQQIGVIMITNLVEKGRRKCDQYWPTESENSEEEYGHILVTLKRTNSLACYTVRHFTIRHMKLRKG (12) RTVIQYHYTQWPDGMT
8. ( 0) EQNTSVIIMITNLVERGRRKCDQYWPDTGTEQYGLITVTLKHTKVLANFTIRTFSLRHAKVKKG ( 3) RTVLQYHYTQWPDHGT
9. ( 567) EQRVAIIVMITNLVERGRRKCDMYWPKDGVETYGVIQVKLIEEEVMSTYTVRTLQIKHLKLKKK ( 6) KLVYQYHYTNWPDHGT
      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

```

```

1. PEYSLPVLTFVRKAAYAKRHA (0) VGPVVVHCSAGVGRTGTIYVLD SMLQQIQHEGTVNIFGFLKHIRSQRNYLVQTEEQYVFIHDTLVE
2. PEYTLPLVTFVRKASHAKRHA (0) VGPIVVHCSAGVGRTGTIYVLD SMLQQIQHEGTVNIFGFLKHIRTQRNYLVQTEEQYIFIHDAIVE
3. PEYALPVLTFVRHSAAKASH (0) TGPVVVHCSAGVGRTGTIYVID SMLQQIKQKGTNVNVLGFLKHIRTQRNYLVQTEEQYIFIHDAIVE
4. PEYALPVLTFVRRSSAARMPE (0) TGPVLVHCSAGVGRTGTIYVID SMLQQIKDKSTNVNVLGFLKHIRTQRNYLVQTEEQYIFIHDAIVE
5. PEYALPVLTFVRRSSAARTPH (0) MGPVVVHCSAGVGRTGTIYVID SMLQQIKDKSTNVNVLGFLKHIRTQRNYLVQTEEQYIFIHDAIVE
6. PEYALPVLTFVRKSSAAQTPD (0) MGPVVVHCSAGVGRTGTIYVLD SMLQKIKDKSTNVNVLGFLKHIRTQRNYLVQTEEQYVFIHDLVLE
7. PEYTLPLVTFVRKSSAARTPE (0) MGPVVVHCSAGVGRTGTIYVLD SMLHQMREQGTANVLGFLKHIRTQRNYLVQTEEQYIFIHEALRE
8. PDYALPVLTFIRKSAAVKVGE (1) VGPVVVHCSAGVGRTGTIYVLD TMMRQVEEKHSVNIKGFLKHIRTQRNYLVQTEEQYIFCHDALLK
9. PDHPLPVLNVFVKSSAANPAE (0) AGPIVVHCSAGVGRTGTIYVLD AMLKQIQQKNIVNVFGLRHIRARQNFVLQTEEQYIFLHDAIVE
      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

```

```

1. AILSKETEVLDSHIHAY ( 9) AGKTKLEKQFQLLSQSNIIQQSDYSALKQCNREKNRT (1) SIIPVERSRVGISSLSG (0) EGTDYINAS
2. AILSKETEVLETHIHAY ( 9) TGKTRLEKQFKLLSQSNTQQCDYSTALKQCNREKNRT (1) SIIPVERSRVGISSLSG (0) EGTDYINAS
3. AILTDEVPANHLHTY ( 9) SGKTKLEKQFKLVSQPHAKQCDYSALKQCNRDKNRN (1) SIIPVERSRVGLATLTA (0) EGSDYINAS
4. AILGKETEVSSNLHSHY ( 9) GGKTRLEKQFKLVTQCNAKYVECFSAQKECNKEKNRN (1) SVVPSERARVGLAPLPG (1) KGTDYINAS
5. AILGKETEVSANQLHSHY ( 9) GGKTRLEKQFKLVTQCNAKYVECFSAQKDCNKEKNRN (1) SVVPSERARVGLAPLPG (1) KGTDYINAS
6. AIHGKETEVHSSQLHGY ( 9) SGKMRLEKQFKLVTQCNAKFIECFSAQKECNKEKNRN (1) SVVPTERARVGLAPLPG (1) KGTDYINAS
7. AVLAGNTEVSCHQFQAY (17) SNSTCLERQFMLVLQMSFRHSECFGAQKHCKNKEKNRN (1) SVVPPVEHARVGLSSLTG (1) EGSDYINAS
8. FIQCRDTEIPQNQLHSHY ( 9) SGRITLDDKHFKHMAHQAEFEYCAQKPCNTKKNRT (1) KLLPVECTRVCLTAKPG (1) EGSDYINAS
9. AIASGETNLMAEQVEEL ( 1) NCTPYLEQYQYKNIQFQPKDHIASAMKQVNSIKNRG (0) AIFPIEGSRVHLTPKPG (1) DGSDYINAS
      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

```

```

1. YIMGYYQSNEFIITQHPLLHTIKDFWRMIWDHNAQLVVMIPDQGN (5) FVYWPKNKDEPINCESFKVTLMAE (5) SNEEKLIQDFILEAT
2. YIMGYYQSNEFIITQHPLLHTIKDFWRMIWDHNAQLVIMLPDSQN (5) FVYWPKNKDEPINCETFKVTMIAE (5) SKEEKLIQDFILEAT
3. YVSGFHRIDEFIVTQHPLPYTTKDFWRMIWDYNAQIIIVMLPDNQG (1) TPYWPKKDGPNNYEAFTVSLLE (5) SNEEQTLIKDFILEAT
4. YIMGYYRSNEFIITQHPLPHTTKDFWRMIWDHNAQIIIVMLPDNQS (5) FVYWPSSREESMNCEAFTVTLISK (5) SNEEQIIIHDFILEAT
5. YIMGYYRSNENVIITQHPLPHTTKDFWRMIWDHNAQIIIVMLPDNQS (5) FVYWPSSREESMNCEAFTVTLISK (5) SNEEQIIIHDFILEAT
6. YIMGYYRSNEFIITQHPLPHTTKDFWRMIWDHNAQIIIVMLPDNHG (5) SVYWPSSREESMNCEAFTVTLISK (5) SNEEQVIIHDFILEAT
7. YIMGYRCSNEFIITQHPLPHTTRDFWRMIWDHNSSETVVMPLPGNQS (5) FEYWPGREEAMTCEFTVTLISK (5) SNEEQVIIHDFILEAT
8. YLQGYNRSKEFIVTQHPLTDTVKDFWQMIWDQNSPAVVMLSKPDE (4) PSYWPSSKDKPMCEVTFKVSQDE (5) GNDISVTVRSFLEAT
9. WLHGFRRLRDFIVTQHPLMAHTIKDFWQMVWDHNAQTVVLLSSLDD (3) AQFWPDEATPIESDHYRVKFLNK (0) TNKSDYVSRDFVQSI
      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

```

```

1. QDDYVLEVRHFQCPKPWNPDSPIKTFELISIVIKKEAANRDGPIMVHDEHGGVTAGTFCALTTLMHQLEKENSVDVYQVAKMINLMRPGV
2. QDDYVLEVRHFQCPKPWNPDSPIKTFELISIIKEETSNRDGPIMVHDEHGGVTAGTFCALTTLMHQLEKENSVDVYQVAKMINLMRPGV
3. QDDYVLEVRHYQSPRPWNPDPGPIKTFELNLIRDAATRDGP IIVHDEHGGITAGTFCALTCLMDQLESENSVDVYQVAKMINLMRPGI
4. QDDYVLEVRHFQCPKPWNPDPAPISSTFELINVIKEEALTRDGPTIVHDEYGAVSAGMLCALTTLSSQLEKENSVDVYQVAKMINLMRPGV
5. QDDYVLEVRHFQCPKPWNPDPAPISSTFELINVIKEEALTRDGPTIVHDEYGAVSAGTLCALTTLSSQLEKENSVDVYQVAKMINLMRPGV
6. QDDYVLEVRHFQCPKPWNPDPAPLSSTFELINVLKEEASTRDGPTIVHDEFGGVSAGLFCGLTTLSSQLEKENSVDVYQVAKMINLMRPGV
7. QDDYVLEVRHFQCPWPNADGPLSRAFELVNLVREASAAHDGPTVHDEFGGVTAGTFCALSTLYQQLLEDQAIQVDFVHAKMINLMRPGV
8. QDDYVLEVRHFQSPDPACTAISDVFHLVDLVQEHAATNEGPITVHDEVYGGVSAATFCGLTALKHQLEDDGSMDFQLAKTYHMMRPGV
9. QDDYELTVKMLHCPSPWPEMSNPNSIYDFIVDVHERCNDYRNGPIVIVDRYGAQACTFCAISSLAIMEVYCSTANVYQYAKLYHNKRPGV
      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

```

```

1. FADIEQYQFLYKVI ( 36)
2. FTDIEQYQFLYKAI ( 36)
3. FTDIDQFQFLYKAI ( 38)
4. FTDIEQYQFIYKAR ( 38)
5. FTDIEQYQFLYKAM ( 38)
6. FTDIEQYQFLYKAM ( 38)
7. FTELEHYHFLYRML ( 35)
8. FSEKDQYFLYKAM ( 29)
9. WTSSEDIRVIYNIL (218)
      *

```

5-6 PTPN6サブタイプのアライメント。番号は図16の分子系統樹上の番号に対応する。

```

1. ( 94) QENSRVIVMTTKEVERGKSKCVKYWPDEMSSKEYGAMRVKNVRENPAHDYILRELKL ( 9) RTVWQYHFKTWPDHGVPSPDPGGVL
2. (346) QENSRVIVMTTKEVERGKSKCVKYWPDEYALKEYGVMVRNVKESAAHDYTLRELKL ( 9) RTVWQYHFKTWPDHGVPSPDPGGVL
3. (346) QENSRVIVMTTKEVERGKSKCVKYWPDEFALKEYGVMVRNVKETPAHDYILRELKL ( 9) RAVWQYHFKTWPDHGVPADPGGVL
4. (346) QENSRVIVMTTKEVERGKSKCVKYWPDEYSLKEYGVMVRNVKESAAHDYTLRELKL ( 9) RTVWQYHFKTWPDHGVPSPDPGGVL
5. ( 2) QEGSRVIVMTTKEVERGKSKCVRYWPELGIEKEFGMLRVGLVDETAKQEYTLRELSL ( 9) RTIWQYQYKAWPDHGVPDGLGEVI
6. ( 89) QENAHVIVMTTKEIERGKCKCVRYWPDVKTSTEYGMISVRNLEERMAQDYIVRDLEV ( 9) RHVWHYQYLSWPDHGVPNEPGGVL
7. (301) QENSRVIVMTTKEVEKGRNKCVPYWPEVGMQRAYGPYSVTNCGEHDTEYKLRTLQV ( 9) REIWHYQYLSWPDHGVPSEPGGVL
8. ( 0) QENTRIIVMTTKEVEKGRTKCEPYWPEEQGKKEFGSYIVQNLEEREATDYKVRVLEV ( 9) RKIHHYQYLSWPDHGVPSEPGGVL
9. ( 2) QENSRVIFMTTKEVERGRSKCVRYWPEEGCSKDYGLRVYNVHESESYDYLRLQLEL ( 8) RIVWQYHFKAWPDHGVPNEPGGVL
10. ( 0) QENTRVIVMTTKEVERGKKNCTRYWPEPGTKKSYGILLIRHLSETNFPDYTLRELEV ( 9) RTIYHYHFKAWPDHGVPDGPDCVL
11. (468) QENTRVIVMTTKEYERGKECARYWPDEGRSEQFGHARIQCVSENSTSDYTLREFLV ( 7) RRIHYHFKAWPDHGVPADPGCVL
12. ( 0) QENSLIIVMTTNEVERGRNKCTRYWPDLDVSKLYGKVTCCVKETTKPHYVLRFEV (14) RTIFQYHFKAWPDHGVPDHPGAVL
      **      * * * * * * * * * *      *      *      *      *      *      *      *      *      *

```

```

1. DFLEEVNIKQENIPEA (0) GPIVVHCSAGIGRTGTFIVIDILIDIIREKGVDCDIDVPKTIQMVRSQRSGMVQTEAQYRFIYMAVQHYIE
2. DFLEEVHHKQESIMDA (0) GPVVVHCSAGIGRTGTFIVIDILIDIIREKGVDCDIDVPKTIQMVRSQRSGMVQTEAQYRFIYMAVQHYIE
3. DFLEEVHHKQDGITDA (0) GPVVVHCSAGIGRTGTFIVIDILIDIVIREKGVDCDIDVPKTIQMVRSQRSGMVQTEAQYRFIYMAVQHYIE
4. DFLEEVHHKQESISDA (0) GPVVVHCSAGIGRTGTFIVIDILIDIIREKGVDCDIDVPKTIQMVRSQRSGMVQTEAQYRFIYMAVQHYIE
5. EFLDEVAKRQAYLPEA (0) GPIVVHCSAGIGRTGTLIVIDILIDIIREKGVDCDIDVPKTIQMVRSQRSGMVQTEAQYRFIYMAVQHYIE
6. SFLDQVNRTQSSIPDT (0) GPIIVHCSAGIGRTGTIIVIDMLISDISRQGLDCDIDIPKTIQIVRKQSRSGMVQTEAQYKFIYMAVQQYIE
7. SFLDQINQRQESLPHA (0) GPIIVHCSAGIGRTGTIIVIDMLMENISTKGLDCDIDIQKTIQMVRAQRSGMVQTEAQYKFIYVAIAQFIE
8. SFLEQVNRRQESMPEA (0) GPIIVHCSAGIGRTGTIIVIDMLVDKINSQGLDCDIDIQKTIQNVRSQRSGMVQTEAQYKFIYMAVQQYIE
9. NFLDEVNRRQDSIPDA (0) GPLVVHCSAGIGRTGTFIVIDILIDIIRHKLDCDIDIPKTIQMVRAQRSGMVQTEAQYKFIYMAVQQYIE
10. NFLHETNTKQENIPDV (0) GPMVVHCSAGIGRTGTFIVIDIILNLIKKFGLDCEIDIQKTIQHMRTQRSGMVQTEAQYKFIYMAVQHHIE
11. NFLQDVNTRQSHLAQA (4) GPICVHCSAGIGRTGTFIVIDMILDQIVRNLGDLTEIDIQRTIQMVRSQRSGMVQTEAQYKFIYAVQHYIQ
12. GILQDVNLQKELTDE (4) GPIVVHCSAGIGRTGTFIVIDVIQDLINHQGDCEIDIQKTIQLVRSQRSGMVQTEQQYKFIYVAIQHVE
      *      *      **      * * * * * * * * * *      *      *      *      *      *      *      *      *

```

```

1. TLQRRRI (11) EYTNIKYS ( 46)
2. TLQRRRI (11) EYTNIKYS ( 45)
3. TLQRRRI (11) EYTNIKYS ( 47)
4. TLQRRRI (11) EYTNIKYS ( 45)
5. NLKKEG (10) EYSNIDAA ( 40)
6. TVQRRRL (11) EYSNIRYP ( 45)
7. TTKKKL (11) EYGNITYP ( 53)
8. TTQKKL (11) EYGNISYL ( 48)
9. TMQQRRL (11) EYTNIRYS ( 47)
10. TVSKRI ( 9) DYTNIKYS ( 75)
11. TLIARK (11) EYTNIKYT (169)
12. SYQAML ( 6) DYGNMTFP ( 23)
      *      *

```


5-7 PTPR2Aサブタイプのアライメント。番号は図17の分子系統樹上の番号に対応する。

```

1. (1444)EQRSATVMMTKLEERSRVKCDQYWPSRGTEETHGLVQVTLTLDVTELATYCVRTFALYKNGSSEKREVRQFQFTAWPDHGVPEHP
2. ( 0)EQRSATVMMTKLEERSRVKCDQYWPSRGTEETHGLVQVTLTLDVTELATYCVRTFALYKNGSSEKREVRQFQFTAWPDHGVPEHP
3. ( 0)EQRSATVMMTKLEERSRVKCDQYWPSRGTEETHGLVQVTLTLDVTELATYCVRTFALYKNGSSEKREVRQFQFTAWPDHGVPEHP
4. (1032)EQRSATVMMTKLEERSRVKCDQYWPSRGTEETHGLVQVTLTLDVTELATYCVRTFALYKNGSSEKREVRQFQFTAWPDHGVPEHP
5. (1480)EQRSATVMMTKLEERSRVKCDQYWPSRGTEETHGLVQVTLTLDVTELATYCVRTFALYKNGSSEKREVRQFQFTAWPDHGVPEHP
6. (1429)EQRTATVMMTKLEERSRVKCDQYWPSRGTEETHGLVQVTLTLDVTELATYCVRTFALYKNGSSEKREVRQFQFTAWPDHGVPEHP
7. ( 0)EQRTATVMMTKLEERSRVKCDQYWPSRGTEETHGLVQVTLTLDVTELATYCVRTFALYKNGSSEKREVRQFQFTAWPDHGVPEHP
8. ( 0)EQRSATVMMTKLEERSRVKCDQYWPSRGTEETHGLVQVTLTLDVTELATYCVRTFALYKNGSSEKREVRQFQFTAWPDHGVPEHP
9. ( 0)EQRSATVMMTKLEERSRVKCDQYWPSRGTEETHGLVQVTLTLDVTELATYCVRTFALYKNGSSEKREVRQFQFTAWPDHGVPEHP
10. ( 0)EQRSATVMMTKLEERSRVKCDQYWPSRGTEETHGLVQVTLTLDVTELATYCVRTFALYKNGSSEKREVRQFQFTAWPDHGVPEHP
11. ( 962)EQRTATVMMTKLEERSRVKCDQYWPSRGTEETHGLVQVTLTLDVTELATYCVRTFALYKNGSSEKREVRQFQFTAWPDHGVPEHP
12. (1576)EQRSATVMMTKLEERSRVKCDQYWPSRGTEETHGLVQVTLTLDVTELATYCVRTFALYKNGSSEKREVRQFQFTAWPDHGVPEHP
13. ( 763)ELKSTIVMMTKLEERSRVKCDQYWPSRGTEETHGLVQVTLTLDVTELATYCVRTFALYKNGSSEKREVRQFQFTAWPDHGVPEHP
14. (1561)ELKSTIVMMTKLEERSRVKCDQYWPSRGTEETHGLVQVTLTLDVTELATYCVRTFALYKNGSSEKREVRQFQFTAWPDHGVPEHP
      *      *      *      *      *      *      *      *      *      *      *      *      *      *      *

```

```

1. TPFLAFLRRVKTCPDAGPMVHCSAGVGRGTGCFIVDAMLERIKHEKTVDIYGHVTLMSQRNVMVQTEDQYIFIHDALEAVTCGNT
2. TPFLAFLRRVKTCPDAGPMVHCSAGVGRGTGCFIVDAMLERIKHEKTVDIYGHVTLMSQRNVMVQTEDQYIFIHDALEAVTCGNT
3. TPFLAFLRRVKTCPDAGPMVHCSAGVGRGTGCFIVDAMLERIKHEKTVDIYGHVTLMSQRNVMVQTEDQYIFIHDALEAVTCGNT
4. TPFLAFLRRVKTCPDAGPMVHCSAGVGRGTGCFIVDAMLERIKHEKTVDIYGHVTLMSQRNVMVQTEDQYIFIHDALEAVTCGNT
5. TPFLAFLRRVKTCPDAGPMVHCSAGVGRGTGCFIVDAMLERIKHEKTVDIYGHVTLMSQRNVMVQTEDQYIFIHDALEAVTCGNT
6. TPFLAFLRRVKTCPDAGPMVHCSAGVGRGTGCFIVDAMLERIKHEKTVDIYGHVTLMSQRNVMVQTEDQYIFIHDALEAVTCGNT
7. TPFLAFLRRVKTCPDAGPMVHCSAGVGRGTGCFIVDAMLERIKHEKTVDIYGHVTLMSQRNVMVQTEDQYIFIHDALEAVTCGNT
8. TPFLAFLRRVKTCPDAGPMVHCSAGVGRGTGCFIVDAMLERIKHEKTVDIYGHVTLMSQRNVMVQTEDQYIFIHDALEAVTCGNT
9. TPFLAFLRRVKTCPDAGPMVHCSAGVGRGTGCFIVDAMLERIKHEKTVDIYGHVTLMSQRNVMVQTEDQYIFIHDALEAVTCGNT
10. TPFLAFLRRVKTCPDAGPMVHCSAGVGRGTGCFIVDAMLERIKHEKTVDIYGHVTLMSQRNVMVQTEDQYIFIHDALEAVTCGNT
11. TPFLAFLRRVKTCPDAGPMVHCSAGVGRGTGCFIVDAMLERIKHEKTVDIYGHVTLMSQRNVMVQTEDQYIFIHDALEAVTCGNT
12. TPFLAFLRRVKTCPDAGPMVHCSAGVGRGTGCFIVDAMLERIKHEKTVDIYGHVTLMSQRNVMVQTEDQYIFIHDALEAVTCGNT
13. TPFLAFLRRVKTCPDAGPMVHCSAGVGRGTGCFIVDAMLERIKHEKTVDIYGHVTLMSQRNVMVQTEDQYIFIHDALEAVTCGNT
14. TPFLAFLRRVKTCPDAGPMVHCSAGVGRGTGCFIVDAMLERIKHEKTVDIYGHVTLMSQRNVMVQTEDQYIFIHDALEAVTCGNT
      *      *      *      *      *      *      *      *      *      *      *      *      *      *      *

```

```

1. EVPARNLAYAIQK ( 9)VTGMELEFKRLASSKAHTSRFISANLPCNFKFNRLVNIIMPYESTRVCLQPIRGVEGSDYINASFIDGYRQQA
2. EVPARNLAYAIQK ( 9)VTGMELEFKRLASSKAHTSRFISANLPCNFKFNRLVNIIMPYESTRVCLQPIRGVEGSDYINASFIDGYRQQA
3. EVPARNLAYAIQK ( 9)VTGMELEFKRLASSKAHTSRFISANLPCNFKFNRLVNIIMPYESTRVCLQPIRGVEGSDYINASFIDGYRQQA
4. EVPARNLAYAIQK ( 9)VTGMELEFKRLASSKAHTSRFISANLPCNFKFNRLVNIIMPYESTRVCLQPIRGVEGSDYINASFIDGYRQQA
5. EVPARNLAYAIQK ( 9)VTGMELEFKRLASSKAHTSRFISANLPCNFKFNRLVNIIMPYESTRVCLQPIRGVEGSDYINASFIDGYRQQA
6. EVPARNLAYAIQK ( 9)VTGMELEFKRLASSKAHTSRFISANLPCNFKFNRLVNIIMPYESTRVCLQPIRGVEGSDYINASFIDGYRQQA
7. EVPARNLAYAIQK ( 9)VTGMELEFKRLASSKAHTSRFISANLPCNFKFNRLVNIIMPYESTRVCLQPIRGVEGSDYINASFIDGYRQQA
8. EVPARNLAYAIQK ( 9)VTGMELEFKRLASSKAHTSRFISANLPCNFKFNRLVNIIMPYESTRVCLQPIRGVEGSDYINASFIDGYRQQA
9. EVPARNLAYAIQK ( 9)VTGMELEFKRLASSKAHTSRFISANLPCNFKFNRLVNIIMPYESTRVCLQPIRGVEGSDYINASFIDGYRQQA
10. EVPARNLAYAIQK ( 9)VTGMELEFKRLASSKAHTSRFISANLPCNFKFNRLVNIIMPYESTRVCLQPIRGVEGSDYINASFIDGYRQQA
11. EVPARNLAYAIQK ( 9)VTGMELEFKRLASSKAHTSRFISANLPCNFKFNRLVNIIMPYESTRVCLQPIRGVEGSDYINASFIDGYRQQA
12. EVPARNLAYAIQK ( 9)VTGMELEFKRLASSKAHTSRFISANLPCNFKFNRLVNIIMPYESTRVCLQPIRGVEGSDYINASFIDGYRQQA
13. EVPARNLAYAIQK ( 9)VTGMELEFKRLASSKAHTSRFISANLPCNFKFNRLVNIIMPYESTRVCLQPIRGVEGSDYINASFIDGYRQQA
14. EVPARNLAYAIQK ( 9)VTGMELEFKRLASSKAHTSRFISANLPCNFKFNRLVNIIMPYESTRVCLQPIRGVEGSDYINASFIDGYRQQA
      *      *      *      *      *      *      *      *      *      *      *      *      *      *      *

```

```

1. YIATQGPLAETTEDFWRLWEHNSTIVVMTKLREMGREKCHQYWPASERSARYQYFVVDPMAYNMPQYILREFKVTDARDGQSRVTRQF
2. YIATQGPLAETTEDFWRLWEHNSTIVVMTKLREMGREKCHQYWPASERSARYQYFVVDPMAYNMPQYILREFKVTDARDGQSRVTRQF
3. YIATQGPLAETTEDFWRLWEHNSTIVVMTKLREMGREKCHQYWPASERSARYQYFVVDPMAYNMPQYILREFKVTDARDGQSRVTRQF
4. YIATQGPLAETTEDFWRLWEHNSTIVVMTKLREMGREKCHQYWPASERSARYQYFVVDPMAYNMPQYILREFKVTDARDGQSRVTRQF
5. YIATQGPLAETTEDFWRLWEHNSTIVVMTKLREMGREKCHQYWPASERSARYQYFVVDPMAYNMPQYILREFKVTDARDGQSRVTRQF
6. YIATQGPLAETTEDFWRLWEHNSTIVVMTKLREMGREKCHQYWPASERSARYQYFVVDPMAYNMPQYILREFKVTDARDGQSRVTRQF
7. YIATQGPLAETTEDFWRLWEHNSTIVVMTKLREMGREKCHQYWPASERSARYQYFVVDPMAYNMPQYILREFKVTDARDGQSRVTRQF
8. YIATQGPLAETTEDFWRLWEHNSTIVVMTKLREMGREKCHQYWPASERSARYQYFVVDPMAYNMPQYILREFKVTDARDGQSRVTRQF
9. YIATQGPLAETTEDFWRLWEHNSTIVVMTKLREMGREKCHQYWPASERSARYQYFVVDPMAYNMPQYILREFKVTDARDGQSRVTRQF
10. YIATQGPLAETTEDFWRLWEHNSTIVVMTKLREMGREKCHQYWPASERSARYQYFVVDPMAYNMPQYILREFKVTDARDGQSRVTRQF
11. YIATQGPLAETTEDFWRLWEHNSTIVVMTKLREMGREKCHQYWPASERSARYQYFVVDPMAYNMPQYILREFKVTDARDGQSRVTRQF
12. YIATQGPLAETTEDFWRLWEHNSTIVVMTKLREMGREKCHQYWPASERSARYQYFVVDPMAYNMPQYILREFKVTDARDGQSRVTRQF
13. YIATQGPLAETTEDFWRLWEHNSTIVVMTKLREMGREKCHQYWPASERSARYQYFVVDPMAYNMPQYILREFKVTDARDGQSRVTRQF
14. YIATQGPLAETTEDFWRLWEHNSTIVVMTKLREMGREKCHQYWPASERSARYQYFVVDPMAYNMPQYILREFKVTDARDGQSRVTRQF
      *      *      *      *      *      *      *      *      *      *      *      *      *      *      *

```

5-7 続き

```

1.  QFTDWPEQGVPKSGEGFIDFIGQVHKTKQFGQDGPISVHCSAGVGRTGVFITLSIVLERMRYEGVVDIFQTVKMLRTQ(1) PAM(0) VQ
2.  QFTDWPEQGVPKSGEGFIDFIGQVHKTKQFGQDGPISVHCSAGVGRTGVFITLSIVLERMRYEGVVDIFQTVKMLRTQ(1) PAM(0) VQ
3.  QFTDWPEQGVPKSGEGFIDFIGQVHKTKQFGQDGPISVHCSAGVGRTGVFITLSIVLERMRYEGVVDIFQTVKMLRTQ(1) PAM(0) VQ
4.  QFTDWPEQGVPKSGEGFIDFIGQVHKTKQFGQDGPISVHCSAGVGRTGVFITLSIVLERMRYEGVVDIFQTVKMLRTQ(0) PAM(0) VQ
5.  QFTDWPEQGVPKSGEGFIDFIGQVHKTKQFGQDGPISVHCSAGVGRTGVFITLSIVLERMRYEGVVDIFQTVKMLRTQ(1) PAM(0) VQ
6.  QFTDWPEQGVPKSGEGFIDFIGQVHKTKQFGQDGPITVHCSAGVGRTGVFITLSIVLERMRYEGVVDVFQTVKMLRTQ(1) PAM(0) VQ
7.  HFIDWPEQGVPKSGEGFIDFIGQVHKTKQFGQDGPITVHCSAGVGRTGVFITLSIVLERMRYEGVVDVFQTVKMLRTQ(1) PAM(0) VQ
8.  QFIDWPEQGVPKSGEGFIDFIGQVHKTKQFGQDGPITVHCSAGVGRTGVFVTLISIVLERMRYEGVVDVFQTVKMLRTQ(1) PAM(0) VQ
9.  QFIDWPEQGTPTGEGFIDFIGQVHKTKQFGQDGPITVHCSAGVGRTGVFVTLISIVLERMRYEGVVDVFQTVKMLRTQ(1) PAL(0) VQ
10. QFTDWPEQGVPKSGEGFIDFIGQVHKTKQFGQEGPITAHCSAGVGRTGVFITLAIVLDRMRYEGVVDVFQTVKMLRTQ(1) PAM(0) VQ
11. QFTDWPEQGVPKSGEGFIDFIGQVHKTKQFGQDGPITVHCSAGVGRTGVFIALSVVLERMRYEGVVDLFQTVRLLRTQ(1) PCC(6) VQ
12. QFTDWPEQGVPKSGEGFIDFIGQVHKTKQFGQEGPITAHCSAGVGRTGVFITLSIVLERMRYEGVVDVFQTVNVLRTQ(1) PGM(0) VQ
13. QFITWPEQGVPKSGQGFIDFIGQVHKTKQFGQDGPITVHCSAGVGRTGVFITLSIVLERMQYEGVLDVFQTVRILRSQ(1) PAM(0) VQ
14. QFIDWPEQGVPKSGEGFIDFIGQVHKTKQFGQDGPITVHCSAGVGRTGVFITLSIVLERMQYEGVLDVFQTVRILRSQ(1) PAM(0) VQ
    *****

```

```

1.  T(0) EDQYQFCYRAALEYLGSFDHYAT(0)
2.  T(0) EDQYQFCYRAALEYLGSFDHYAT(0)
3.  T(0) EDEYQFCYQAALYLGSFDHYAT(0)
4.  T(0) EDEYQFCYQAALYLGSFDHYAT(0)
5.  T(0) EDEYQFCYQAALYLGSFDHYAT(0)
6.  T(0) EDQYQLCYRAALEYLGSFDHYAT(0)
7.  T(0) EDQYQLCYRAALEYLGSFDHYAT(0)
8.  T(0) EDQYQFCYRAALEYLGSFDHYAT(0)
9.  I(0) EEQYQFCYRAALEYLGSFDHYAV(0)
10. T(0) EDQYQFCYRAALEYLGSFDHYAN(1)
11. T(1) EDHYAFCYRAALEYLGSFRHYAN(0)
12. T(0) EEQYAFYRAALEYLGSFDHYTN(0)
13. T(0) EDQYQFCYRAALEYLGSFDHYAT(0)
14. T(0) EDQYHFCYRAALEYLGSFDNYTN(0)
    * * *

```


5-8 PTPR3サブタイプのアライメント。番号は図19の分子系統樹上の番号に対応する。

```

1. ( 739) EQKATVIVMVT RCEEGNRNKCAEYWP (8) FGDVVVKINQHKRCPDYIIQKLN (11) VTHIQFT
2. ( 329) EQNTATIVMVTNLKERKECKCAQYWP (6) YGNIRVSVEDVTVLVDYTVRKFCI (14) ITQFHFT
3. ( 223) EQKSATIVMLTNLKERKEEKCHQYWP (6) YGNIRVCVEDCVVLVDYTVRKFCI (13) VSQFHFT
4. ( 996) QENSASIVMVTNLVEVGRVKCVRYWP (5) YGDIKVTLIETEP LAEYVIRFTTV (10) LRLFHFT
5. ( 987) HENTASIIMVTNLVEVGRVKCKYWP (5) YKDIKVTLIETEL LAEYVIRTFV (10) IRQFHFT
6. ( 974) QEQSACIVMVTNLVEVGRVKCYKYWP (5) YGDFKVTCEMEPLAEYVVRTFTL (10) VKQFHFT
7. ( 967) QEHCSIVMITKLVEVGRVKCSRYWP (5) YGDIKIMLVKTETLA EYVVRTFAL (10) VRQFHFT
8. (1561) ELKTATIVMMTRLEERTRIKCDQYWP (6) YGQIFVTITETQELATYSIRTFQL (10) IKQLQFT
9. (1444) EQRSATVMMTRLEERSRVKCDQYWP (6) HGLVQVTLTLDTVELATYCVRTFAL (10) VRQFQFT
10. (1480) EQRSATVMMTRLEEKSRVKCDQYWP (6) YGFIQVTLTLDTIELATFCVRTFSL (10) VRQFQFT
11. (1429) EQRTATVMMTRLEEKSRVKCDQYWP (6) CGLIQVTLTLDTVELATYTVRTFAL (10) LRQFQFM
12. ( 0) EGQLQTIVMLTKCSEGGKIKCEMYWS (8) TDTLTVTTTTSNESYADFQIRKFSI (12) VTHFHYT
13. (1128) EKNVYAIIMLT KCVEQGRTKCEEYWP (6) YGDITVAMTSEIVLPEWTIRDFTV (10) LRQFQFT
14. (1197) EKNYISIVMLTKCVEQARTKCEQYWP (6) YGDIIVTMVSEVVLPEWTIRDFNV (10) VRQFHFT
15. (1791) EQNVHNIVMVTQCVEKGRVKCDHYWP (7) YGDLILQMLSESVLPEWTIREFKI (11) IRHFHYT
16. ( 2) EQGVHVVMLTQCVERGRVKCDHYWP (7) YGDLVVQMLSESVLPEWTLREFKV (11) VRQFHYT
17. (1026) QQKSQIIVMLTQCNEKRRVKCDHYWP (7) YGDITVEMISEEEQDDWACRHFRI ( 8) VMHFNYT
18. (1077) QQKSQIIVMLTQCNEKRRVKCDHYWP (7) YGDITVEMLSEEEHTDWVYRNFRI ( 8) VMHFNYT
19. (1360) ESNSRAIVMLTRCFEKGREKCDQYWP (7) YGDIKVQILNDSHYADWVMTFEML ( 8) LRHFHFT
20. (1417) ESNSRAIVMLTRCFEKGREKCDQYWP (7) YGDIKVQLIIDTHYHDSISEFMV ( 8) MRHFHFT
21. ( 0) EQNSRVIVMLTQCWERGKPKCERYWP (7) YGDIVVKMLSENKEDDWTREFEL ( 8) VRQYQFT
22. (2094) ETRTKTLVMLTQCFEKGRIRCHQYWP (8) FGDIIVTKLMEDIQIDWTIRDLKI ( 8) VRQCNFT
23. (1238) EQQVHVIIIMLTVMGEMNVRVLCHEYWP (7) HGHITHTLLAESEDEWTRREFQL (10) VKQLQFT
24. ( 910) EQQSHTLVMLTNCMEAGRVKCEHYWP (7) HGHLRVTLVGEEVMENWTVRELL (10) VRQFHYQ

```

* * * * *

```

1. SWPDHGVP (0) EDPHLLKLRRRVN (7) GPIVVHCSAGVGRGTGTIIGIDAMLEGLEAENKVDVYGVVVKL
2. SWPDEGVP (0) FTPIGMLKFLKKVK (7) GAIIVHCSAGVGRGTGTFFVIDAMLDMHHTERKVDVYGFVSRI
3. SWPDEGVP (0) FTPIGMLKFLKKVK (7) GPIVVHCSAGVGRGTGTFFVIDAMMAMHAEQVDVFEFVSRI
4. SWPDHGVP (0) CYATGLLGFVRQVK (7) GPIVVHCSAGAGRTGCFIAIDTMLDMAENEGVVDIFNCVREL
5. GWPDEGVP (0) YHATGLLGFVRQVK (7) GPLVVHCSAGAGRTGCFIVIDIMLMAEREGVVDIYNVREL
6. GWPDEGVP (0) YHATGLLSFIRRVK (7) GPIVVHCSAGAGRTGCFIVIDIMLMAEREGVVDIYNVREL
7. AWPDEGVP (0) YHATGLLAFIRRVK (7) GPIVHCSAGTGTGTGCFIVLDVMLDMAECEGVVDIYNVCKTL
8. AWPDEGVP (0) DHPAPFLQFLRRCR (7) GPVIVHCSAGVGRGTGCFIVIDSMLERMKHEKIIDIYGHVTCL
9. AWPDEGVP (0) EHPFPFLAFLRRVK (7) GPMVVHCSAGVGRGTGCFIVIDAMLERIKHEKTVDIYGHVTCL
10. AWPDEGVP (0) EYPTPFLAFLRRVK (7) GPIVVHCSAGVGRGTGCFIVIDAMLERIKPEKTVDIYGHVTCL
11. AWPDEGVP (0) EYPTPFLAFLRRVK (7) GPMVVHCSAGVGRGTGCFIVIDAMLERIKHEKTVDIYGHVTCL
12. AWPDEGVP (0) QTATSILSFVRRVQ (8) VPLLHCSAGVGRGTGTFFIALDRLDRVRSSETISVFEIVKDM
13. SWPDHGVP (0) DTDLLINFRYLVR (9) SPILVHCSAGVGRGTGTFFIAIDRLIYQIENENTVDVYGVVYDL
14. SWPDHGVP (0) ETDLLINFRHLVH (9) SPILVHCSAGVGRGTGTFFIAIDRLIYQIENENTVDVYGVVYDL
15. VWPDEGVP (0) ETTQSLIQFVRTVR (9) GPTVVHCSAGVGRGTGTFFIALDRILQQLDSKDSVDIYGAVHDL
16. VWPDEGVP (0) EMAGSLIKFVRTVQ (9) GPTIVHCSAGVGRGTGTFFIALDRILHQMADAASWVDIHGTVMDL
17. AWPDEGVP (2) NAAESILQFVHMVR (7) GPMIIVHCSAGVGRGTGTFFIALDRLLQHIRDHEFVDILGLVSEM
18. AWPDEGVP (2) NAAESILQFVQMV (7) GPMIIVHCSAGVGRGTGTFFIALDWLQHIRDHEFVDILGLVSDM
19. TWPDEGVP (0) NPPQTLVRVRAFR (7) RPIIVHCSAGVGRSGTFTLDRILQQLINTSDYVDIFGIVYAM
20. TWPDEGVP (0) EPPLSLVRVRAFR (7) RPIIVHCSAGVGRSGTFTLDRILQQLIHKSVDYVDIFGIVFAM
21. SWPDHGVP (0) EDTSSLQFVHMVR (7) GPTVVHCSAGVGRSGTFTLDRLLQHMDHEQVDIFGIVHQM
22. GWPDEGVP (0) ENTPLIHFKVLVR (7) TPMVVHCSAGVGRGTGFIALDHLTQHINNHDVFVDIYGLVAEL
23. TWPDSV (0) EAPSSLLAFVELVQ (9) GPILVHCSAGVGRGTGTFFVALLPAVRQLEEEQVVDVFNFTVYIL
24. AWPDEGVP (0) SSPDTLLAFWRMLR (9) GPPIVHCSAGVGRGTGTFFIALDVLLRQLQSEGLLGPFSFVRKM

```

** **

***** ** *

*

5-8 続き

- 1. RRQRCLMVQVEAQYILIHQAL (397)
 - 2. RAQRCQMVQTD MQYVFIYQAL (304)
 - 3. RNQRPQMVQTD MQYTFTYQAL (309)
 - 4. RAQRVNLVQTEEQYV FVHDAI (301)
 - 5. RSRRVNMVQTEEQYV FFIHDAI (301)
 - 6. RSRRINMVQTEEQYI FFIHDAI (301)
 - 7. CSRRVNMVQTEEQYI FFIHDAI (305)
 - 8. RAQRNYMVQTEEQYI FFIHDAI (303)
 - 9. RAQRNYMVQTEEQYI FFIHDAL (303)
 - 10. RSQRNYMVQTEEQYS FIHEAL (303)
 - 11. RSQRNYMVQTEEQYV FIHEAL (303)
 - 12. RNRRRFMIQTLAQYV LIYDAF (13)
 - 13. RMHRPLMVQTEEQYV FLNQCV (42)
 - 14. RMHRPLMVQTEEQYV FLNQCV (42)
 - 15. RLHRVHVMVQTECQYV YLHQCV (37)
 - 16. RSHRVLMVQTECQYA FLHQCV (36)
 - 17. RSYRMSMVQTEEQYI FFIHQCV (24)
 - 18. RSYRMSMVQTEEQYI FFIHQCV (24)
 - 19. RKERVWMVQTEQQYI CIHQCL (107)
 - 20. RKERVFMVQTEQQYV CIHQCL (186)
 - 21. RMHRVFMVQTQEYI FFIHQCI (53)
 - 22. RSERMCMVQNLAQYI FLHQCI (43)
 - 23. RLHRPLMIQTL SQYIFLHSCL (299)
 - 24. RESRPLMVQTEAQYV FLHQCI (40)
- * * **

5-9 PP遺伝子族のアライメント。番号は図22の分子系統樹上の番号に対応する。

```

1. ( ) DLELLIDDAAMMPKERMVVE ( 3 ) ITVVGDVHADIEALRKIER ( 5 ) A (1) FLGDYADRGDY (0) PVECYEKVLR (1) FLEG (
2. ( 3 ) ELPRLLEDADVEVLRGGPVVRL ( 4 ) VLVAGDTHGYPEVSRWVLG ( 9 ) V (1) FLGDYVDRGPR (0) GVENLSLLVE (1) LLAE (
3. ( 4 ) ETYELLEKSKEIDFRQGGPPIGE ( 5 ) VVFGDTHGAINVTEYVFR ( 8 ) I (1) FLGDYVDRGPQ (1) GVENLSLLK (0) KLIE (
4. ( 14 ) ELLLLLPEIDRIFDSEPAVLRI ( 4 ) VMVGDHGNLQALEYVIE ( 9 ) F (1) FLGDYVDRGPQ (0) GTEALIRLFR (1) KLED (
5. ( 10 ) IPSALEDIGPIDMGSSVVPV ( 3 ) LVVVGDTHTGFEVIGKILA (11) Y (1) FNGDYVDRGSF (0) SIEITVLLA (1) KIAR (
6. ( 14 ) HLHYQTKKGAGVQVELNIDRN ( 3 ) IAVIGDLHGDAESTKYIFQ (13) A (1) FLGDYVDRGQH (0) GINVLTSILA (1) KVVY (
7. (228) VMRELVRAGELMKNSTCLAL ( 7 ) LAIVGDHIGSLPDLCAINE (12) A (1) FLGDYVDRGPK (0) GHTVVTALLC (1) KLCE (
8. (123) YVALILREAAKSLQLPNISPV ( 7 ) VTVCGDLHGKLDLDDLVVLH (10) Y (1) FNGDYVDRGKR (0) GLEVLLSLLS (1) YLAF (
9. (144) YVNLILYETKHLVQLPNINRV ( 7 ) ITVCGDLHGKLDLDTIFIFY (10) Y (1) FNGDYVDRGKD (0) SVEILMILFA (1) MLVY (
10. (137) YVLEVLFTKVKLKQMPNFTHI ( 7 ) VTCIGDLHGKLDLDFLIFY (10) Y (1) FNGDYVDRGKN (0) SIEILMILCV (1) FLVY (
11. (169) AAWRVFTDAMSHLNTMPNVRL (18) VVVVGDLHGQLADLLHLK (10) Y (1) FNGDYVDRGAN (0) GVEVLLILFS (1) MLAC (
12. ( 48 ) VFDLSLVTAKHILRQKNCVHI ( 8 ) VVVVGDLHGQLHDLFLK (10) Y (1) FNGDYVDRGAW (0) GLETFVLVLS (1) KVLN (
13. (212) YVAARI SHADTLFRQEPMSVEL ( 9 ) ISVCGDTHGQFYDVLNLF (10) Y (1) FNGDYVDRGSW (0) SCEVALLFYC (1) KILH (
14. (182) FAYSILRDLKELLEKTPLSDI ( 7 ) LVICGDTHGQFYDILLNIFK (10) Y (1) FNGDYVDRGSW (0) STEVAFTLYA (1) KLLY (
15. (186) YVYQIIIVARNIVNEFTMVEV ( 7 ) LTVCGDTHGQFYDLMELFR (10) Y (1) FNGDYVDRGSW (0) STEIALLLYA (1) KWLK (
16. (199) CAYQILVQKVEVLSKSLTLET ( 7 ) ITVCGDTHGQFYDILLNIFE (10) Y (1) FNGDYVDRGSF (0) SVEVILTLEF (1) KLLY (
17. (242) YAYQIVLQTRQIILLALPSLVDI ( 7 ) ITVCGDVHGQFYDILLNIFE (10) Y (1) FNGDYVDRGSF (0) SVEIILTLFA (1) KCMC (
18. ( 59 ) VALRIITEGASILRQKELLNDI ( 3 ) VTVCGDIHGQFFDLMLKFE ( 9 ) Y (1) FLGDYVDRGYF (0) SIECVLYLWA (1) KILY (
19. (133) AALRIIQEGATLRTKEKTMIDI ( 3 ) VTVCGDIHGQFYDLMKLF ( 9 ) Y (1) FLGDYVDRGYF (0) SIECVLYLWS (1) KITY (
20. ( 54 ) VALRIITEGAALRREKXNDIV ( 3 ) ITVCGDIHGQFFDLVLF ( 9 ) Y (1) FLGDYVDRGYF (0) SIECVLYLWS (1) KITY (
21. ( 68 ) IALRIINEGAAILRREKTMIEV ( 3 ) ITVCGDIHGQFFDLMLKFE ( 9 ) Y (1) FLGDYVDRGYF (0) SIECVLYLWV (1) KILY (
22. ( 55 ) VALRIINDGAAILRQKNTMIEV ( 3 ) ITVCGDIHGQFFDLMLKFE ( 9 ) Y (1) FLGDYVDRGYF (0) SIECVLYLWS (1) KINH (
23. ( 57 ) QALWIITREGTKLRAEPNMLE ( 3 ) ITVCGDVHGQFYDLMKLF ( 9 ) Y (1) FLGDYVDRGYF (0) SIECVLYLWA (1) KIHV (
24. ( 89 ) QALYILEKGGVILMEESNIQPV ( 3 ) VTVCGDIHGQFYDLMKLF ( 9 ) Y (1) FLGDYVDRGYF (0) SIECVLYLWS (1) KMWY (
25. ( 88 ) QTLVYIKKATILKESNDLIEV ( 3 ) VTVCGDIHGQFYDLMKFE ( 9 ) Y (1) FLGDYVDRGYF (0) SIECVLYLWA (1) KIWY (
26. ( 88 ) QALIKNMSTVALSKPEPNLLK ( 3 ) ITICGDIHGQFYDLMKLF ( 9 ) Y (1) FLGDYVDRGAF (0) SFECLLYLYS (1) KLNK (
27. (113) QAARIVTLATELSKEPNLISV ( 3 ) ITVCGDIHGQFYDLMKLF ( 9 ) Y (1) FLGDYVDRGSF (0) SFECLLYLYS (1) KLNK (
28. (107) DVEIVKMAAIELEKEPTLIQV ( 3 ) ITVCGDTHGQFYDLIKIFE (11) Y (1) FLGDYVDRGYF (0) SMEVITYLYA (1) KINY (
29. ( 56 ) DLYKLVGDCNKLILMEESNVLY ( 3 ) LTVVGDIHGQFYDLMKLE (11) Y (1) FLGDYVDRGSY (0) SIEVITLLYA (1) KLNK (
30. ( 91 ) DIRTIVERGCAVLNAQPNVLYI ( 3 ) IYVGDTHGQLHDLILNIFY ( 9 ) Y (1) FLGDYVDRGYQ (0) SVECLLLMLV (1) KIKR (
31. (20) QLKRLCEMVKILLMEESNIQPV ( 3 ) VTVCGDIHGQFYDLMELFR ( 9 ) Y (1) FNGDYVDRGYF (0) SLETFTLFML (1) KARY (
32. (22) EMKQCEMVKILLMEESNIQPV ( 3 ) VTVCGDIHGQFHDLMELFR (10) Y (1) FLGDYVDRGY (0) SLETFTLFMC (1) KVKY (
33. (22) DLKRLCDYVCDLLEESNVQPV ( 3 ) VTVCGDIHGQFYDLCELFR ( 9 ) Y (1) FNGDYVDRGY (0) SLETFTYLLA (1) KAKW (
34. (20) ELKRLCEMVCDLLEETNIPV ( 3 ) VTVCGDIHGQFYDLQELFR ( 9 ) Y (1) FNGDYVDRGY (0) SLETFTRLLT (1) KARY (
35. (19) ELQLLCEVYKILLMEESNVQPV ( 3 ) VTVCGDIHGQFHDLMELFQ ( 9 ) Y (1) FNGDYVDRGYN (0) SLEVFTLLLL (1) KARH (
36. (23) DLKRLCEVYKILLMEESNVQPV ( 3 ) VTVCGDIHGQFYDLMELFR ( 9 ) Y (1) FNGDYVDRGY (0) SLETFTYLLA (1) KARY (
37. (24) DLRLVQCRVKEILVEENVSQI ( 3 ) VVICGDIHGQFFDLLELFD ( 9 ) Y (1) FLGDYVDRGYN (0) SVETFEYLLL (1) KLLF (
38. (22) NVVLEQCRVRIILMEESNVQI ( 3 ) VVICGDIHGQFHDLMELFR ( 9 ) Y (1) FLGDYVDRGY (0) SVETFLLLLT (1) KCKY (
39. (20) EVKALCLKAMEILVEESNVQV ( 3 ) VVICGDIHGQFYDMKELFK ( 9 ) Y (1) FLGDYVDRGY (0) SVETFLLLLA (1) KVRY (
40. (21) EVKALCSKAREILVEESNVQV ( 3 ) VVICGDIHGQFYDLKELFK ( 9 ) Y (1) FNGDYVDRGY (0) SVETFLLLLA (1) KVRY (
41. (23) EVKALCAKAREILVEESNVQV ( 3 ) VTVCGDIHGQFYDLKELFR ( 9 ) Y (1) FNGDYVDRGY (0) SVETFLLLLA (1) KVRY (
42. (23) EVKALCAKAREILVEESNVQV ( 3 ) VTVCGDIHGQFYDLKELFK ( 9 ) Y (1) FNGDYVDRGY (0) SVETFLLLLA (1) KVRY (
43. (20) DVKILCNKAKDILNEENNVIRV ( 3 ) VVICGDIHGQFYDLMEVFK ( 9 ) Y (1) FLGDYVDRGYN (0) SVETFLLLLA (1) KVRY (
44. (20) TVERLCLNSQELLMNEGNVQV ( 3 ) VVICGDIHGQLHDLTLFE ( 9 ) Y (1) FLGDYVDRGY (0) SLESFLLLLC (1) KLRY (
45. (19) TVRALCFKLKEMLVKESNVHI ( 3 ) VTVVGDHMQFHDMLIEFQ ( 9 ) Y (1) FLGDYVDRGLY (0) SVETIMLLIV (1) KLRY (
46. (20) VIAYLCLKAEVILMEESNVQV ( 3 ) ITVVGDIHGQFDDLELIEFR ( 9 ) Y (1) FLGDYVDRGY (0) SIETITLLIC (1) KLRY (
47. (177) EVKLMCDLILGILKNEENCVR ( 3 ) VTVAGDIHGQFFDLLELFI ( 9 ) Y (1) FLGDYVDRGY (0) SCECFCLVAC (1) KIKY (
48. (26) QVKSLEKAKELILTESNVQV ( 3 ) VTVCGDVHGQFHDLMELFR ( 9 ) Y (1) FNGDYVDRGY (0) SVETVTLVLA (1) KVRY (
49. (26) QVRTLCEKAKELILTESNVQV ( 3 ) VTVCGDVHGQFHDLMELFR ( 9 ) Y (1) FNGDYVDRGY (0) SVETVTLVLA (1) KVRY (
50. (26) QVRTLCEKAKELILTESNVQV ( 3 ) VTVCGDVHGQFHDLMELFR ( 9 ) Y (1) FNGDYVDRGY (0) SVETVTLVLA (1) KVRY (
51. (23) EVRDLCEKAREILSKESNVQV ( 3 ) VTVCGDIHGQFHDLMELFK ( 9 ) Y (1) FNGDYVDRGY (0) SVETVTLVLA (1) KVRY (
52. (44) DVQRLEKAREVILQDESNVQV ( 3 ) VTVCGDIHGQFHDLMELFK ( 9 ) Y (1) FNGDYVDRGY (0) SVETVTLVLA (1) KIRY (
53. (23) NVKELVAKAREVFSKKNVQV ( 3 ) VTVCGDIHGQFHDLMELFK ( 9 ) Y (1) FNGDYVDRGYN (0) SVETVTLVLS (1) KVRY (
54. (30) QVRLCEKAREILMEESNVQV ( 3 ) VVICGDIHGQFHDLAELFR ( 9 ) Y (1) FNGDYVDRGY (0) SVETVTLVVG (1) KVRY (
55. (29) DVQICEKAREILMEESNVQV ( 3 ) VVICGDIHGQFHDLMELFK ( 9 ) Y (1) FNGDYVDRGH (0) SVECVTLVTL (1) KIRH (
56. (39) DVEMLCCKAREVILQENNVQV ( 3 ) VTVCGDIHGQFHDLMELFK ( 9 ) Y (1) FNGDYVDRGY (0) SVETVSLVLA (1) KLRY (
57. (26) DVIMQCDLKAKEVLSVESNVQV ( 3 ) VTVCGDIHGQFHDLMELFN ( 9 ) Y (1) FNGDYVDRGY (0) SVETVSLLIA (1) KIRY (
58. (23) DVKILCDQAKAILVEENNVQV ( 3 ) VTVCGDIHGQFYDLIELFR ( 9 ) Y (1) FNGDYVDRGY (0) SVETVSLVLA (1) KIRY (
59. (94) DVARLCKMAVDVLFEENVPK ( 3 ) VVICGDIHGQFHDLMELFK ( 9 ) Y (1) FNGDYVDRGY (0) SVETVSLVLA (1) KVRY (
60. (69) EIREICERVAHIFIEENPIQ ( 3 ) VTVVGDIHGQYHDLIEMFR ( 9 ) Y (1) FNGDYVDRGY (0) SVECTITLVL (1) KIKY (
61. (20) EVMKLCNKARSILDEGENTHVI ( 3 ) CTIVGDIHAQFFDLIEMFR ( 9 ) Y (1) FNGDYVDRGH (0) GVETFLLLLA (1) KVKY (
62. (27) EIVEGTHRAEILQEPNVTLI ( 3 ) ITLVGDIHGQFYDLREMF ( 9 ) Y (1) FLGDYVDRGYF (0) SVETFLLLLA (1) KVYF (
63. (31) EVRSLCIKAKEILQDGPPIQV ( 3 ) LTICGDIHGQFHDLIELFQ ( 9 ) Y (1) FLGDYVDRGSQ (0) SVETFSLLMS (1) KVRY (
64. (68) DVIALLLWRAELLSSEPTVPI ( 5 ) LNVCGDVHGQFYDVRYMFE (10) Y (1) FLGDYVDRGY (0) SLEITLILLC (1) KVNK (
65. (66) DVIILVWRVQAILQEPNVTLI ( 3 ) VNICGDIHGQFYDVVKLF (43) Y (1) FLGDYVDRGYF (0) SLEITLILYL (1) KIMY (
66. (415) EVKHLCEDEVEKIFMNEPTLQL ( 3 ) KIVFGDIHGQYDLMRLFI (15) Y (1) FLGDYVDRGQH (0) SLEIMLLFA (1) KIEY (
67. (407) EIAADLCSAERIFSEPTVLQL ( 3 ) KIFGDLHGQFYDLMRLFI (15) Y (1) FLGDYVDRGQH (0) SLEITLILLA (1) KVEY (
68. (557) EVELCYAAEQIFMHEQTVLQL ( 3 ) KIVFGDLHGQFYDLMRLFI (15) Y (1) FLGDYVDRGQH (0) SLEITLILLA (1) KIEY (
69. (551) NISVLCISIVIDIFKQEDMVLK ( 3 ) IKIYGDHIGQYDLMRMFQ (24) Y (1) FLGDYVDRGSN (0) SLEVICLLFA (1) KCKY (
70. (270) EILQICHTAREVILQEPNVTLI ( 3 ) IKVVGDVHGQFYDMLRLILK ( 9 ) Y (1) FLGDYVDRGKN (0) SLEITLILLC (1) KIKY (
71. (388) EILQICHTAREVILQEPNVTLI ( 3 ) KIVVGDVHGQYDMLRLIFT ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILFC (1) KIKY (
72. (423) EILQICHTAREVILQEPNVTLI ( 3 ) KIVVGDVHGQYADMLRLIFT ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLC (1) KIKY (
73. (233) EIVAICHAREVILQEPNVTLI ( 3 ) KIVVGDVHGQYDMLRMFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILMC (1) KIKY (
74. (217) EITSICMAVREIFLQEPNVTLI ( 3 ) KIVVGDVHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILFL (1) KIRY (
75. (46) EIEAVCSRAREVLLKQPTLLET ( 3 ) INLLGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KARY (
76. (29) LIERLIQTRREVILQEPNVTLI ( 3 ) VNICGDIHGQFYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILFA (1) KVKY (
77. (71) EVRVLVMSRAREVILQEPNVTLI ( 3 ) VVICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KLSF (
78. (39) EIRQLCIVSREIFLQEPNVTLI ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KIKY (
79. (46) EIRQLCAVSREIFLQEPNVTLI ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KIKY (
80. (43) EIKQLCTTARDIFLQEPNVTLI ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KIKY (
81. (35) EIRQLCFNARQIFLQEPNVTLI ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KIRY (
82. (29) EIKQLCFVSARDIFLQEPNVTLI ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KIKF (
83. (30) EIKHLCTAKQIFLQEPNVTLI ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KIKY (
84. (30) EIRQLCLASREIFLQEPNVTLI ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KVKY (
85. (30) EIRQLCLTSKEILSEQPNLLEL ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KIKY (
86. (30) EIRQLCLTAKDVMSQPNLLEL ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KVKY (
87. (30) EIRQLCLTAKEIFMSQPNLLEL ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLS (1) KIKY (
88. (29) EIRQLCTTSRSIFLQEPNVTLI ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEVICLLFA (1) KIKY (
89. (32) EVRGLCLKSREIFLQEPNVTLI ( 3 ) LIICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KIKY (
90. (32) EVRGLCLKSREIFLQEPNVTLI ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KIKY (
91. (31) EIRSLCLKSREIFLQEPNVTLI ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KIKY (
92. (33) EIRGLCLKSREIFLQEPNVTLI ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KIKY (
93. (35) EVRGLCLKARDIFLQEPNVTLI ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KIKY (
94. (30) EICALCIRSRDIFMDQEPNVTLI ( 3 ) KLVCGDVHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KVKY (
95. (29) EIRYLTVQATREIFLQEPNVTLI ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KIKY (
96. (32) EIRFLCNKAREIFLQEPNVTLI ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEVICLLLA (1) KIKY (
97. (32) EIRYLCSKARSIFLQEPNVTLI ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KIKY (
98. (33) EIRYLCTKAREIFLQEPNVTLI ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEITLILLA (1) KIKY (
99. (42) EIEELVNKAEIFISHPVILLEL ( 3 ) IKICGDIHGQYDMLRLIFE ( 9 ) Y (1) FLGDYVDRGKQ (0) SLEICITLLC (1) KIKY (

```


1. 0) KALLLRGNHSTGV (0) YPHQLPYQLR (2) F (6) VYESLIRMWNKMPLS (0) AIV (1) G (1) IWMHGGVP (5) IHAE (1) I (1) R
 2. 2) RVLLRGNHESPSM (1) LYGFGEFE (2) V (3) HLDILHGFFYSCLPYA (0) ALM (0) G (1) VLLVHGGVP (13) LGEI (1) E (1) L
 3. 8) KIVVLRGNHESPLT (1) PHYGFEEELK (2) T (5) AYEKFRDLFSYMPYS (0) VLV (1) G (0) YLCMHGGLP (12) VEDI (1) K (1) K
 4. 2) HIFLLRGNHETVDM (1) IYGFPAEIG (0) F (3) FLERVSTQYDKMPIA (0) AVL (1) G (1) TFCVHGGIN (4) INDI (1) K (0) E
 5. 2) AVHLLRGNHETESV (1) SNYSLPAEIE (2) Y (1) LLAMNKAQALPLC (0) AVI (1) D (1) YLVVHGGIS (6) LQDM (1) K (1) N
 6. 2) RVITIRGNHESNL (1) RRYGFDDEV (2) Y (4) YHKTIAFPFKLLPVA (8) TTL (1) N (1) ILFVHGGLP (4) LDTL (0) I (1) R
 7. 2) RVFLLRGNHETVSM (1) SFFGYRQVD (2) Y (5) MFDMTSLFATMPLC (0) CLI (1) K (1) IFCTHGGAP (6) GDEI (1) K (1) I
 8. 2) AVFLNRGNHESVM (1) ARYGFTRVE (2) Y (6) ILAFIDEVYRWLPLG (0) SVL (1) S (1) VLVHGGFS (4) LDLI (1) S (1) D
 9. 2) EFHLNRGNHEDHMV (1) LRYGFTKEVM (2) Y (6) ILRTLDQVFCWLPLA (0) TLI (1) E (1) VLLHGGVS (4) LELL (1) K (1) E
 10. 2) DLHLNRGNHEDFMM (1) RYGFTEKIL (2) Y (6) ILQILEEFYAWLPIG (0) TIV (1) N (1) ILVIHGGIS (4) LNL (1) R (1) E
 11. 2) VYTLNRGNHEDCYM (1) DEYGFDEVS (2) Y (3) VFRLVQRCFCALPLA (0) TII (1) K (1) VVVHGGLP (6) IEDI (1) R (1) Q
 12. 2) RVYLLRGNHESKYC (1) SMYGFKEVL (2) Y (6) VYRKCLCGFEGLPLA (0) SII (1) G (1) VYTAHGGLP (32) LDEL (1) Q (1) R
 13. 2) NFFLLRGNHESDNN (1) KIYGFDECK (2) Y (3) IFNMFQSFESLPLA (0) TLI (1) N (1) YLVHGGLP (6) LSDF (1) N (1) D
 14. 2) AVFNLNRGNHETDDM (1) KVGFECECR (2) Y (3) TFI FSETFSLLPLG (0) SLI (1) D (1) YLVHGGLP (6) LDQL (1) N (1) D
 15. 2) GFFNLNRGNHETDDM (1) RVYGFGECK (2) Y (3) TYKLFSESFSALPLA (0) TLI (1) K (1) FLVHGGLP (6) LDDI (1) K (1) D
 16. 2) HFHLLRGNHETDNN (1) QIYGFGEVK (2) Y (3) MYELFSEVFEWLPLA (0) QCI (1) G (1) VLMHGGLP (6) LDDI (1) K (1) E
 17. 2) SIYLARGNHESKSM (1) KIYGFGEVR (2) L (3) FVDLFAEVFCYLPLA (0) HVI (1) G (1) VVVHGGLP (6) LDDI (1) A (1) D
 18. 2) TLFLRGNHESRHL (1) EYFTFKQEC (2) Y (3) VYDACMDAFDCLPLA (0) ALM (1) Q (1) FLCVHGGLS (5) LDDI (1) K (1) D
 19. 2) TLFLRGNHESRHL (1) EYFTFKQEC (2) Y (3) VYDACMDAFDCLPLA (0) ALM (1) Q (1) FLCVHGGLS (5) LDDI (1) R (1) D
 20. 2) TLFLRGNHESRHL (1) EYFTFKQEC (2) Y (3) VYDACMDAFDCLPLA (0) ALL (1) Q (1) FLCVHGGLS (5) LDDI (1) T (1) N
 21. 2) TLFLRGNHESRHL (1) EYFTFKQEC (2) Y (3) VYDACMDAFDCLPLA (0) ALL (1) Q (1) FLCVHGGLS (5) LDDI (1) R (1) D
 22. 2) TLFLRGNHESRHL (1) EYFTFKQEC (2) Y (3) VYDACMDAFDCLPLA (0) ALL (1) Q (1) FLCVHGGLS (5) LDDI (1) K (1) D
 23. 2) TLFLRGNHESRHL (1) EYFTFKQEC (2) Y (3) VYDACMDAFDCLPLA (0) ALL (1) Q (1) FLCVHGGLS (5) LDDI (1) N (1) D
 24. 2) TLFLRGNHESRHL (1) EYFTFKQEC (2) Y (3) VYDACMDAFDCLPLA (0) ALL (1) Q (1) FLCVHGGLS (5) LDDI (1) S (1) N
 25. 2) TLFLRGNHESRHL (1) EYFTFKQEC (2) Y (3) VYDACMDAFDCLPLA (0) ALL (1) Q (1) FLCVHGGLS (5) LDDI (1) M (1) N
 26. 2) RFWMRGNHESRHL (1) SYFTFKQEC (2) Y (3) VYDACMDAFDCLPLA (0) ALL (1) Q (1) FLCVHGGLS (5) LDDI (1) K (1) N
 27. 2) HFWMRGNHESRHL (1) SYFTFKQEC (2) Y (3) VYDACMDAFDCLPLA (0) ALL (1) Q (1) FLCVHGGLS (5) LDDI (1) N (1) N
 28. 2) TFFLLRGNHESRHL (1) EYFTFKQEC (2) Y (3) VYDACMDAFDCLPLA (0) ALL (1) Q (1) FLCVHGGLS (5) LDDI (1) N (1) D
 29. 2) TVYFLRGNHESRHL (1) AFNFKEEC (2) Y (3) TYDMLMDSFDLPLA (0) CII (1) S (1) FIAHGGIS (5) LDDI (1) R (1) D
 30. 2) NFFIRGNHETDL (4) RGASLKEEC (2) Y (6) VFNILMRAMDSLPLV (0) AII (1) R (1) ILVHGGIS (4) LEKL (1) E (0) Q
 31. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) I (1) A
 32. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) V (1) S
 33. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) T (1) E
 34. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) T (1) D
 35. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) L (1) E
 36. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) I (1) S
 37. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) L (1) N
 38. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) L (1) D
 39. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 40. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 41. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) S (1) D
 42. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) I (1) N
 43. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) T (1) D
 44. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) I (1) D
 45. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) V (1) D
 46. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) K (1) D
 47. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 48. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 49. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 50. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 51. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 52. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 53. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 54. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 55. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 56. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 57. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 58. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 59. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 60. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 61. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 62. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 63. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 64. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 65. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 66. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 67. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 68. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 69. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 70. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 71. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 72. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 73. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 74. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 75. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 76. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 77. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 78. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 79. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 80. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 81. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 82. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 83. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 84. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 85. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 86. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 87. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 88. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 89. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 90. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 91. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 92. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 93. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 94. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 95. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 96. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 97. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 98. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D
 99. 2) KITLLRGNHESRQI (1) QVGFYDECQ (2) Y (4) VWKYCCQVDFDLPLA (0) AVI (1) N (1) ILVHGGIS (5) LDQI (1) A (1) D

```

1. KE (11) IMWNDP (11) RGVG (0) YFF (1) KKNSEMLLKELVGVKVIIRAHEP (0) Q (1) VLKAEQD (2) VVTIGSCALPY (32)
2. RG (17) LLWSDP (11) RGGG (1) YYY (1) RLAWKGFLEANGLELIVRAHEA (17) E (1) GYTMGLE (5) VVTVSSSLYHG (25)
3. YP (11) LLWNDP (13) RCEG (1) YFF (1) ERVTNEFLSYNNFKGIIRGHEA (1) D (0) GFRTNMN (2) VITVFSVYHG (32)
4. KA (2) YLWNDP (10) RGST (1) KEF (1) PDIVDGLKTNHLKIRVGRHTA (1) D (1) GYRWTFD (2) LLSLFSFCDYV (26)
5. RF (19) ILWADP (12) RGTS (0) VIF (1) AVAAQRFLRTSELKYIIRSHC (1) D (1) GWQVEYP (1) VFTLFSVPNYA (74)
6. SK (29) ILWSDP (17) RGAG (0) ILY (1) SEFYRKATENRVLAMFRGHQV (1) K (1) GFRDFF (1) HYTVFSSSDYV (58)
7. PV (12) TLWSDP (20) RGVG (0) LVY (1) YRAFQWAKKHGKIKLFRHEA (2) P (1) GVRFDFF (8) HYTVFSSSNYV (24)
8. RG (26) IMWSDP (11) RGAG (0) WVF (1) EDVTDNLFQRHRLSVVIRSHC (1) P (1) GHEFMHD (2) IITIFSASNY (276)
9. RS (133) ILWSDP (11) RGGG (0) CYF (1) PDVTQQLLQKYNLQELIRSHC (1) P (1) GYEFCHN (2) VLTIFSASNY (240)
10. RN (55) ILWSDP (11) RGGG (0) CYF (1) PDVTSKILWYQLKMLIRSHC (1) P (1) GYEICHD (2) VVTIFSASNY (225)
11. RF (19) LLWSDP (11) RGAG (0) VVF (1) ADVTQEFLLQNNGLVIRSHC (1) L (1) GYEEHHD (2) LLTVFSASNYD (270)
12. RS (15) VLWSDP (11) RGIG (0) LLW (1) PDCTEDFLKYEKLIRSHC (13) D (1) GYTDHND (5) LITIFSAPDYP (70)
13. RF (11) LLWADP (10) RGLG (0) HAF (1) PDITDRFLRNNKLRKIIRSHC (1) M (1) GVQFEQK (2) LMTVFSAPNYC (54)
14. RF (13) MLWTD (10) RGVG (0) LQF (1) PDVSKRFCEANGLKAVIRSHC (1) D (1) GYEVEHD (2) CITVFSAPNYC (44)
15. RH (13) MLWTD (10) RGVG (0) MQF (1) PDVTKRFCDKNGLEAIRSHC (1) M (1) GYEEHHD (2) CITVFSAPNYC (46)
16. RN (11) LLWSDP (10) RGVG (0) CQF (1) PDVTKAFLENNLDYIRSHC (1) A (1) GYEVHHD (2) CVTVFSAPNYC (47)
17. RF (11) LLWSDP (10) RGVG (0) LSF (1) GDVTKRFLQNNLDLIRSHC (1) D (1) GYEVEHD (2) LITVFSAPNYC (46)
18. RF (11) LLWSDP (10) RGGG (0) YFF (1) YPAVCEFLQNNLDLIRSHC (1) D (1) GYRMYRK (8) LITIFSAPNYL (209)
19. RF (11) LLWSDP (10) RGGG (0) YFF (1) YPAVCEFLQNNLDLIRSHC (1) D (1) GYRMYRK (8) LITIFSAPNYL (192)
20. RF (11) LLWSDP (10) RGGG (0) YFF (1) YPAVCEFLQNNLDLIRSHC (1) D (1) GYRMYRK (8) LITIFSAPNYL (270)
21. RF (11) LLWSDP (10) RGGG (0) YFF (1) YPAVCEFLQNNLDLIRSHC (1) D (1) GYRMYRK (8) LITIFSAPNYL (203)
22. RF (11) LLWSDP (10) RGGG (0) YFF (1) YPAVCEFLQNNLDLIRSHC (1) D (1) GYRMYRK (8) LITIFSAPNYL (194)
23. RF (11) LLWADP (10) RGGG (0) YFF (1) YPAVCEFLQNNLDLIRSHC (1) D (1) GYRMYRK (8) LITIFSAPNYL (208)
24. RF (11) LLWADP (10) RGGG (0) YFF (1) YPAVCEFLQNNLDLIRSHC (1) D (1) GYRMYRK (8) LITIFSAPNYL (297)
25. RF (11) LLWSDP (10) RGGG (0) YFF (1) YPAVCEFLQNNLDLIRSHC (1) D (1) GYRMYRK (8) LITIFSAPNYL (213)
26. RF (11) LLWADP (25) RGGG (0) YAF (1) FKASCKFLKANGLLSIRSHC (1) D (1) GYRMYRK (8) LITIFSAPNYL (205)
27. RF (11) LLWADP (42) RGGG (0) YAF (1) YRAACHFLQETGLLSIRSHC (1) D (1) GYRMYRK (8) LITIFSAPNYL (214)
28. RF (11) LLWSDP (17) RGGG (0) YLY (1) YRAACHFLQNNLDLIRSHC (1) N (1) GYKMHQ (8) VITLFSAPNYL (262)
29. RF (11) LLWSDP (17) RGGG (0) WFF (1) NEAASKFLQNNLDLIRSHC (1) L (1) GYKMHQ (8) VITLFSAPNYC (199)
30. KL (17) LLWSDP (30) RGGG (0) EKT (1) DELSCSILDKYGFDLVRGHEC (1) H (1) GMSSTVS (2) VFTVFGATHYH (33)
31. RA (11) LMWSDP (10) RGAG (0) WLF (1) SKVTTEFSQINDLTLIRAHQL (1) Q (1) GYKYHFA (3) LVTVWSAPNYC (44)
32. RA (11) LMWSDP (10) RGAG (0) WLF (1) SKVAREFHNHNLIRAHQL (1) M (1) GFKYHFA (3) VITVWSAPNYC (47)
33. RN (11) LVWSDP (10) RGAG (0) WLF (1) AKVTNEFHNHNLIRAHQL (1) H (1) GYKMHQ (2) LVTVWSAPNYC (43)
34. RN (11) LVWSDP (10) RGAG (0) WLF (1) HNVTKDFMAINNLIRAHQL (1) N (1) GYKMHQ (2) LVTVWSAPNYC (43)
35. RN (11) LMWSDP (10) RGAG (0) WLF (1) SRVTTEFHNHNLIRAHQL (1) Q (1) GLKYMFQ (3) LVTVWSAPNYC (43)
36. RN (11) LMWSDP (10) RGAG (0) WLF (1) SKVTTEFHNHNLIRAHQL (1) Q (1) GYRMYRK (2) LVTVWSAPNYC (42)
37. RV (11) LMWSDP (10) RGAG (0) WLF (1) PNVTKFHNHNLIRAHQL (1) M (1) GYRMYRK (3) IITVWSAPNYC (43)
38. RK (11) LMWSDP (10) RGAG (0) FLF (1) ADVSEVFNHNDLIRAHQL (1) M (1) GYKMHQ (44) VITVWSAPNYC (44)
39. RK (11) LMWSDP (11) RGAG (0) FLF (1) GSVVTSFHNHNLIRAHQL (1) M (1) GYKMHQ (2) IITVWSAPNYC (44)
40. RK (11) LMWSDP (10) RGAG (0) FLF (1) EDVVQKFNHNDLIRAHQL (1) M (1) GYKMHQ (2) LVTVWSAPNYC (44)
41. RK (11) LMWSDP (10) RGAG (0) YLF (1) SDVVAQFNAANDLIRAHQL (1) M (1) GYKMHQ (2) LVTVWSAPNYC (44)
42. RK (11) LMWSDP (10) RGAG (0) YLF (1) SDVVSQFNHNDLIRAHQL (1) M (1) GYKMHQ (2) LVTVWSAPNYC (44)
43. RK (11) LMWSDP (10) RGAG (0) FVF (1) ADVVKEFNHNDLIRAHQL (1) M (1) GYKMHQ (2) LVTVWSAPNYC (43)
44. RK (11) LMWSDP (10) RGAG (0) FVF (1) KREVDQFLEKNNLIRAHQL (1) M (1) GYKMHQ (2) LVTVWSAPNYC (48)
45. RF (11) LVWSDP (22) RGAG (0) YTF (1) RSVVKEFLRNMNRIYRAHQL (1) N (1) GYQIYFD (2) VITVWSAPNYC (96)
46. RF (11) LVWSDP (11) RGAG (0) FVF (1) EVTVTKFLENNHNLIRAHQL (1) S (1) GYQIYFD (2) LSTVWSAPNYC (46)
47. RF (11) LVWSDP (15) RGAG (0) YLF (1) EETKTESFLRNMNLSIRAHQL (1) Q (1) GFQNMHN (2) VVTIFSAPNYC (43)
48. RL (11) LVWSDP (10) RGAG (0) YTF (1) QDISETFNHANGLTLVSRHQL (1) M (1) GYNWCHD (2) VVTIFSAPNYC (43)
49. RL (11) LVWSDP (10) RGAG (0) YTF (1) QDISETFNHANGLTLVSRHQL (1) M (1) GYNWCHD (2) VVTIFSAPNYC (43)
50. RL (11) LVWSDP (10) RGAG (0) YTF (1) QDISETFNHANGLTLVSRHQL (1) M (1) GYNWCHD (2) VVTIFSAPNYC (43)
51. RV (11) LVWSDP (10) RGAG (0) YTF (1) KDISETFNHANGLTLVSRHQL (1) M (1) GYNWCHD (2) VVTIFSAPNYC (43)
52. RI (11) LVWSDP (10) RGAG (0) YTF (1) QDISETFNHANGLTLVSRHQL (1) M (1) GYNWCHD (2) VVTIFSAPNYC (43)
53. RI (11) LVWSDP (10) RGAG (0) YTF (1) QDISETFNHANGLTLVSRHQL (1) M (1) GYNWCHD (2) VVTIFSAPNYC (44)
54. RV (11) LVWSDP (10) RGAG (0) YTF (1) QDISETFNHANGLTLVSRHQL (1) M (1) GYNWCHD (2) VVTIFSAPNYC (43)
55. RV (11) LVWSDP (10) RGAG (0) YTF (1) QDISETFNHANGLTLVSRHQL (1) M (1) GYNWCHD (2) VVTIFSAPNYC (43)
56. RV (11) LVWSDP (10) RGAG (0) YTF (1) QDISETFNHANGLTLVSRHQL (1) M (1) GYNWCHD (2) VVTIFSAPNYC (43)
57. RV (11) LVWSDP (10) RGAG (0) YTF (1) QDISETFNHANGLTLVSRHQL (1) M (1) GYNWCHD (2) VVTIFSAPNYC (43)
58. RI (11) LVWSDP (10) RGAG (0) YTF (1) QDISETFNHANGLTLVSRHQL (1) M (1) GYNWCHD (2) VVTIFSAPNYC (43)
59. RI (11) LVWSDP (10) RGAG (0) YTF (1) QDISETFNHANGLTLVSRHQL (1) M (1) GYNWCHD (2) VVTIFSAPNYC (43)
60. RV (11) LVWSDP (10) RGAG (0) YTF (1) QDISETFNHANGLTLVSRHQL (1) M (1) GYNWCHD (2) VVTIFSAPNYC (43)
61. RR (11) LVWSDP (10) RGAG (0) YTF (1) QDISETFNHANGLTLVSRHQL (1) M (1) GYNWCHD (2) VVTIFSAPNYC (43)
62. RN (11) TLWSDP (13) RGAG (0) FVF (1) SDVVQEFCAANSIELVSRHQL (1) M (1) GYKMHQ (2) LITLFSAPNYC (44)
63. RV (11) LVWSDP (11) RGAG (0) WTW (1) CDITEKFLHNSNKLIRAHQL (1) M (1) GYKMHQ (2) LITLFSAPNYC (43)
64. RS (11) LQWSDP (11) RGAG (0) YLF (1) PRAVDQFLRINDLIRAHQL (1) I (1) GYQYFEN (2) VVTIFSAPNYC (35)
65. RS (11) LQWSDP (11) RGAG (0) FVF (1) AQAVDKFNSDNGVEIIRAHQL (1) P (1) GYQYFEN (2) CCTWSCSPNYC (44)
66. RP (12) LLWSDP (13) RGGG (1) VSF (1) PDIVKAFLENNLIRAHQL (1) I (1) GYQYFEN (2) LITLFSAPNYC (61)
67. RP (12) LLWSDP (13) RGGG (1) VTF (1) PDRVMEFCNNNDLIRAHQL (1) M (1) GYQYFEN (2) LITLFSAPNYC (54)
68. RP (12) LLWSDP (13) RGGG (1) VTF (1) PDRVMEFCNNNDLIRAHQL (1) M (1) GYQYFEN (2) LITLFSAPNYC (74)
69. RP (12) LLWSDP (13) RGGG (1) VTF (1) PDRVMEFCNNNDLIRAHQL (1) M (1) GYQYFEN (2) LITLFSAPNYC (49)
70. RP (11) LLWSDP (11) RGVG (0) YCY (1) KRVNLDKFAKFDLIRAHQL (1) E (1) GYEFFAR (2) FVTIFSAPNYC (39)
71. RP (11) LLWSDP (11) RGVG (0) YCY (1) KRVNLDKFAKFDLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (47)
72. RP (11) LLWSDP (11) RGVG (0) YCY (1) KRVNLDKFAKFDLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (47)
73. RP (11) LLWSDP (11) RGVG (0) YCY (1) KRVNLDKFAKFDLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (58)
74. RP (11) LLWSDP (11) RGVG (0) YCY (1) KRVNLDKFAKFDLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (58)
75. RP (11) LLWSDP (11) RGVG (0) YCY (1) KRVNLDKFAKFDLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (38)
76. RP (11) LLWADL (11) RGVG (0) FTF (1) KRVNLDKFAKFDLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (45)
77. RP (11) LLWSDP (12) RGVG (0) FTF (1) ESUVKSFNKKFDLIRAHQL (1) D (1) GYEFFAR (2) LVTIFSAPNYC (34)
78. RP (11) LLWSDP (11) RGVG (0) FTF (1) PDKVAEFLKNDMLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (33)
79. RP (11) LLWSDP (10) RGVG (0) FTF (1) ADKVAEFLKNDMLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (33)
80. RP (11) LLWSDP (11) RGVG (0) FTF (1) PDKVSEFLKNDMLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (38)
81. RP (11) LLWSDP (11) RGVG (0) FTF (1) ADKVAEFLKNDMLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (43)
82. RP (11) LLWSDP (11) RGVG (0) FTF (1) SDIVSGFLKNDMLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (48)
83. RP (11) LLWSDP (11) RGVG (0) FTF (1) ADKVAEFLKNDMLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (52)
84. RP (11) LLWADP (11) RGVG (0) FTF (1) ADKVAEFLKNDMLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (52)
85. RP (11) LLWADP (11) RGVG (0) FTF (1) PDTVTEFLQKHDLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (49)
86. RP (11) LLWSDP (11) RGVG (0) FTF (1) PDTVTEFLQKHDLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (49)
87. RP (11) LLWADP (11) RGVG (0) FTF (1) PDTVTEFLQKHDLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (34)
88. RP (11) LLWSDP (11) RGVG (0) FTF (1) ADVVSRFLQKHDLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (53)
89. RP (11) LLWSDP (11) RGVG (0) FTF (1) VDVVSFLKNDMLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (58)
90. RP (11) LLWSDP (11) RGVG (0) FTF (1) VDVVSFLKNDMLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (55)
91. RP (11) LLWSDP (11) RGVG (0) FTF (1) AEVVAKFLKNDMLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (56)
92. RP (11) LLWSDP (11) RGVG (0) FTF (1) AEVVAKFLKNDMLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (50)
93. RP (11) LLWSDP (11) RGVG (0) YVF (1) QEVVQFLKNDMLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (45)
94. RP (11) LLWSDP (11) RGVG (0) FTF (1) PDKVAEFLKNDMLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (32)
95. RP (11) LLWADP (11) RGVG (0) FTF (1) ADVVSRFLKNDMLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (52)
96. RP (11) LLWSDP (11) RGVG (0) FTF (1) PDKVAEFLKNDMLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (55)
97. RP (11) LLWSDP (11) RGVG (0) FTF (1) PDKVAEFLKNDMLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (40)
98. RP (11) LLWSDP (11) RGVG (0) FTF (1) PDKVAEFLKNDMLIRAHQL (1) E (1) GYEFFAR (2) LVTIFSAPNYC (35)
99. RP (11) LLWSDP (11) RGVG (0) FTF (1) HDIVMKFLKNDMLIRAHQL (1) Q (1) GYEFFAR (2) LVTIFSAPNYC (56)

```


5-10 PP2Aサブタイプのアライメント。番号は図28の分子系統樹上の番号に対応する。

```

1. (90) YYSVETVTLV LALKVRYRERITILRGNHESRQITQVYGFYDECLRKYGNANVWKYFTDLFDYLP L TALVDGQIFCLHGGLSPSIDT
2. (90) YYSVETVTLV LALKVRYRERITILRGNHESRQITQVYGFYDECLRKYGNANVWKYFTDLFDYLP L TALVDGQIFCLHGGLSPSIDT
3. ( 2) YYSVETVTLV LALKVRYRERITILRGNHESRQITQVYGFYDECLRKYGNANVWKYFTDLFDYLP L TALVDGQIFCLHGGLSPSIDT
4. ( 0) YYSVETVTLV LALKVRYRERITILRGNHESRQITQVYGFYDECLRKYGNANVWKYFTDLFDYLP L TALVDGQIFCLHGGLSPSIDT
5. ( 0) YYSVETVTLV LALKVRYRERITILRGNHESRQITQVYGFYDECLRKYGNANVWKYFTDLFDYLP L TALVDGQIFCLHGGLSPSIDT
6. (90) YYSVETVTLV LALKVRYPERITILRGNHESRQITQVYGFYDECLRKYGNANVWKYFTDLFDYLP L TALVDGQIFCLHGGLSPSIDT
7. (90) YYSVETVTLV LALKVRYPERITILRGNHESRQITQVYGFYDECLRKYGNANVWKYFTDLFDYLP L TALVDGQIFCLHGGLSPSIDT
8. (90) YYSVETVTLV LALKVRYPERITILRGNHESRQITQVYGFYDECLRKYGNANVWKYFTDLFDYLP L TALVDGQIFCLHGGLSPSIDT
9. ( 2) YYSVETVTLV LALKVRFKPRITILRGNHESRQITQVYGFYDECLRKYGNANVWKYFTDLFDYLP L TALVDGQIFCLHGGLSPSIDT
10. ( 0) YYSVETVSLV LALKVRFPNRVITILRGNHESRQITQVYGFYDECLRKYGNPNVWKYFTDLFDYLP L TALVDGQIFCLHGGLSPSIDT
      *****  **  ****  *  *****  *****  *****  *****

```

```

1. LDHIRALDRLQEV PHEGPMCDLLWSDPDDRGGWGISPRGAGYTFGQDISETFNHANG L TLV SRAHQLVMEGYNWCHDRNVVTIF SAPNYC
2. LDHIRALDRLQEV PHEGPMCDLLWSDPDDRGGWGISPRGAGYTFGQDISETFNHANG L TLV SRAHQLVMEGYNWCHDRNVVTIF SAPNYC
3. LDHIRALDRLQEV PHEGPMCDLLWSDPDDRGGWGISPRGAGYTFGQDISETFNHANG L TLV SRAHQLVMEGYNWCHDRNVVTIF SAPNYC
4. LDHIRALDRLQEV PHEGPMCDLLWSDPDDRGGWGISPRGAGYTFGQDISETFNHANG L TLV SRAHQLVMEGYNWCHDRNVVTIF SAPNYC
5. LDHIRALDRLQEV PHEGPMCDLLWSDPDDRGGWGISPRGAGYTFGQDISETFNHANG L TLV SRAHQLVMEGYNWCHDRNVVTIF SAPNYC
6. LDHIRALDRLQEV PHEGPMCDLLWSDPDDRGGWGISPRGAGYTFGQDISETFNHANG L TLV SRAHQLVMEGYNWCHDRNVVTIF SAPNYC
7. LDHIRALDRLQEV PHEGPMCDLLWSDPDDRGGWGISPRGAGYTFGQDISETFNHANG L TLV SRAHQLVMEGYNWCHDRNVVTIF SAPNYC
8. LDHIRALDRLQEV PHEGPMCDLLWSDPDDRGGWGISPRGAGYTFGQDISETFNHANG L TLV SRAHQLVMEGYNWCHDRNVVTIF SAPNYC
9. LDHIRALDRLQEV PHEGPMCDLLWSDPDDRGGWGISPRGAGYTFGQDISETFNHANG L TLV SRAHQLVMEGYNWCHDRNVVTIF SAPNYC
10. LDHIRSLDRLQEV PHEGPMCDLLWSDPDDRGGWGISPRGAGYTFGQDISETFNHNTNGLTLI SRAHQLVMEGYNWCHERNVVTIF SAPNYC
      *****  *****  *****  *****  *****  *****  *****  *****

```

```

1. YRCGNQAAIMELDDTLKYSFLQFDPAPRRGEPHVTRRTPDYFL
2. YRCGNQAAIMELDDTLKYSFLQFDPAPRRGEPHVTRRTPDYFL
3. YRCGNQAAIMELDDTLKYSFLQFDPAPRRGEPHVTRRTPDYFL
4. YRCGNQAAIMELDDTLKYSFLQFDPAPRRGEPHVTRRTPDYFL
5. YRCGNQAAIMELDDTLKYSFLQFDPAPRRGEPHVTRRTPDYFL
6. YRCGNQAAIMELDDTLKYSFLQFDPAPRRGEPHVTRRTPDYFL
7. YRCGNQAAIMELDDTLKYSFLQFDPAPRRGEPHVTRRTPDYFL
8. YRCGNQAAIMELDDTLKYSFLQFDPAPRRGEPHVTRRTPDYFL
9. YRCGNQAAIMELDDTLKYSFLQFDPAPRRGEPHVTRRTPDYFL
10. YRCGNQAAIMELDDGLKYTFLQFDPAPRRGEPLVTRRIPDYFM
      *****  *****  ****  *****  *****  *****

```

5-11 PP2Bサブタイプのアライメント。番号は図29の分子系統樹上の番号に対応する。

```

1. (123) YFSIECVLYLWALKILYPKTLFLLRGNHECRHLTEYFTFKQECKIKYSERVYDACMDAFDCLPLAALMNQQFLCVHGGLSPEINT
2. (123) YFSIECVLYLWALKILYPKTLFLLRGNHECRHPTEYFTFKQECKIKYSERIYDACMDAFDCLPLAALMNQQFLCVHGGLSPEINT
3. ( 0) YFSIECVLYLWALKILYPKTLFLLRGNHECRHLTEYFTFKQECKIKYSERIYDACMDAFDCLPLAALMNQQFLCVHGGLSPEVNT
4. (119) YFSIECVLYLWLSLKINHPKTLFLLRGNHECRHLTDYFTFKQECKIKYSEQVYDACMETFDCLPLAALLNQQFLCVHGGMSPEITS
5. ( 0) YFSIECVLYLWLSLKILYPKTLFLLRGNHECRHLTEYFTFKQECKIKYSEQIYDACMEAFDCLPLAALMNQQFLCVHGGLSPEIHT
6. ( 0) YFSIECVLYLWALKILYPKTLFLLRGNHECRHLTEYFTFKQECKIKYTEHVYDACMEAFDCLPLAALMNQQFLCVHGGLSPEVHN
7. (132) YFSIECVLYLWVLKILYPSTLFLLRGNHECRHLTEYFTFKQECKIKYSERVYACMEAFDCLPLAALLNQQFLCVHGGLSPEIHT
8. ( 0) YFSIECVLYLWALKMLYPKTLFLLRGNHECRHLTEYFTFKQECKIKYKEEVYDACMDAFDCLPLAALMNQQFLCVHGGLSPEIHS
9. (197) YFSIECVLYLWLSLKITYPQTLFLLRGNHECRHLTEYFTFKQECKIKYSERVYDACMDAFDCLPLAALMNQQFLCVHGGLSPEIHE
10. (118) YFSIECVLYLWLSLKITYPTTSLLRGNHECRHLTEYFTFKQECKIKYSESIYDACMEAFDCLPLAALLNQQFLCIHGGLSPEIFT
*****

```

```

1. LDDIRKLDRFKEPPAYGPMCDILWSDPLEDFGNEKTQEHEFTHTNVRGCSYFYSPAVCEFLQHNLLSILRAHEAQDAGYRMYRKSQTTG
2. LDDIRKLDRFKEPPAYGPMCDILWSDPLEDFGNEKSQEHEFTHTNVRGCSYFYSPAVCEFLQHNLLSVLRAHEAQDAGYRMYRKSQTTT
3. LEDIKKLDRFKEPPAYGPMCDLLWSDPLEDFGNEKSQEHEFTHTNVRGCSYFYSPAVCEFLQHNLLSVIRAHEAQDAGYRMYRKSQTTG
4. LDDIRKLDRFTEPPAFGPVCDLLWSDPSEDYGNKLTLEHYHTNVRGCSYFYSPAVCEFLQNNLLSIRAHEAQDAGYRMYRKSQATG
5. LDDIKKLDRFKEPPAFGPMCDLLWSDPLEDFGNEKTQEYFSHTNVRGCSYFYSPAVCEFLQNNLLSVIRAHEAQDAGYRMYRKSQTTG
6. LDDIRKLDRFKEPPAFGPMCDLLWSDPSEDYGNKLTQEYFSHTNVRGCSYFYSPAVCEFLQNNLLSIRAHEAQDAGYRMYRKSQTTG
7. LDDIRKLDRFKEPPAFGPMCDLLWSDPSEDYGNKLTQEYFSHTNVRGCSYFYSPAVCEFLQNNLLSIRAHEAQDAGYRMYRKSQTTG
8. LDDIRKLDRFKEPPAFGPMCDLLWSDPLEDFGNEKNAEHFSHNSVRGCSYFYSPAVCEFLQNNLLSIRAHEAQDAGYRMYRKSQATG
9. LEDIRKLDRFKEPPAFGPMCDLLWSDPLEDFGNEKNSDYTHNSVRGCSYFYSPAVCEFLQNNLLSIRAHEAQDAGYRMYRKSQTTG
10. LDDIKTLNRFREPPAYGPMCDLLWSDPLEDFGNEKTNEFFSHNSVRGCSYFYSPACCEFLQNNLLSIVRAHEAQDAGYRMYRKNQVTG
*****

```

```

1. FPSLITIFSAPNYLDVYNNKAAVLKYENNVNIRQFNCSHPHYWLPNFMDVFTWSLPFVGEKVTEMLVNVNLNICSDELGS (12) RKEVI
2. FRSLITIFWAANYLDVYNNKAAVLKYENNVNIRQFNCSAHPYWLPNFIDVFTWSLPFVGEKVTEMLVNVLSICSDELVS (12) RKEVI
3. FPSLITIFSAPNYLDVYNNKAAVLKYENNVNIRQFNCSHPHYWLPNFMDVFTWSLPFVGEKVTEMLVNVNLNICSDELIS (13) RKEVI
4. FPSLITIFSAPNYLDVYNNKAAVLKYENNVNIRQFNCSHPHYWLPNFMDVFTWSLPFVGEKVTEMLVNVNLNICSDELIS (10) RKEII
5. FPSLITIFSAPNYLDVYNNKAAVLKYENNVNIRQFNCSHPHYWLPNFMDVFTWSLPFVGEKVTEMLVNVNLNICSDELMT (13) RKEVI
6. FPSLITIFSAPNYLDVYNNKAAVLKYENNVNIRQFNCSHPHYWLPNFMDVFTWSLPFVGEKVTEMLVNVNLNICSDELVT (13) RKEVI
7. FPSLITIFSAPNYLDVYNNKAAVLKYENNVNIRQFNCSHPHYWLPNFMDVFTWSLPFVGEKVTEMLVNVLSICSDELMT (12) RKEII
8. FPSLITIFSAPNYLDVYNNKAAVLKYENNVNIRQFNCSHPHYWLPNFMDVFTWSLPFVGEKVTEMLVNVNLNICSDELMT (21) RKEII
9. FPSLITIFSAPNYLDVYNNKAAVLKYENNVNIRQFNCSHPHYWLPNFMDVFTWSLPFVGEKVTEMLVNVNLNICSDELMT (14) RKEVI
10. FPSLITIFSAPNYLDVYNNKAAVLKYENNVNIRQFNCSHPHYWLPNFMDVFTWSLPFVGEKVTEMLVNLNICSDELVA (40) RKEII
*****

```

```

1. RNKIRAIGKMARVFSVLREESESVLTLKGLTPTGMLPSGVLGGKQTL (12) AIKGFSPQHKITSFEEAKGLDRINERMPPRR (34)
2. RNKIRAIGKMARVFSVLREESESVLTLKGLTPTGMLPSGVLGGKQTL (12) AIKGFSPQHKITSFEEAKGLDRINERMPPRR (31)
3. RNKIRAIGKMAKMFVSLREESESVLTLKGLTPTGMLPSGVLGGKQTL (12) AIRGFSPQHKISSFAEAKGLDRINERMPPRR (29)
4. RNKIRAIGKMARVFSILRQESSESVLTLKGLTPTGTLPLGVLGGKQTL ( 2) AIRGFSLQHKIRSFEEAKGLDRINERMPPRR (31)
5. RNKIRAIGKMARVFSVLREESESVLTLKGLTPTGMLPSGVLGGKQTL (18) AIRGFTPQHKISSFEEAKGLDRINERMPPRR (30)
6. RNKIRAIGKMAKMFVSLREESESVLTLKGLTPTGMLPSGVLGGKQTL (12) AIRGFSPQHKITSFEEAKGLDRINERMPPRR (74)
7. RNKIRAIGKMARVFSVLREESESVLTLKGLTPTGMLPSGVLGGKQTL (12) AIRGFSPPHRCSFEEAKGLDRINERMPPRR (28)
8. RNKIRAIGKMARVFTVLREESESVLHLKGLTPTGQLPLGALSGGKQTL ( 2) AIQGYSVNTKIQSFEEAKSLDRINERMPPRR (20)
9. RNKIRAIGKMARVFSVLREESESVLTLKGLTPTGALPLGALSGGKQSL ( 2) AMQGFSPNHKITSFAEAKGLDAVNERMPPRR (25)
10. RNKIRAIGKMSRVFSILREESESVLTLKGLTPTGALPVGALSGGRDSL ( 2) ALQGLTASSHIHSFAEAKGLDAVNERMPPRR (77)
*****

```

5-12 PP1サブタイプのアライメント。番号は図30の分子系統樹上の番号に対応する。

```

1. (103) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRFSVRLWKVFTDSFNCLPVAIVIDDKILCMHGGLSPDLTN
2. (103) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRFNVKLWKVFTDTFNCLPVAIVIDDKILCMHGGLSPELIN
3. (107) KQSLETICLLLAYKIKYPGNFFLLRGNHECASINRIYGFYDECKRRFNVVWVWVFTDCFNCLPVAALIDDKILCMHGGLSPDLHD
4. ( 94) KQSVETICLLLAYKIKYKFNFFLLRGNHECASINRIYGFYDECKRRYSVRVWKIFTDCFNCLPVAALIDDKILCMHGGLSPELKH
5. ( 94) KQSIETICLLLAYKVYKFNFFLLRGNHECASINRVYGFYDECKRRYNVRLWKTFTTECFNCLPVSAALIDDKILCMHGGLSPDIKS
6. ( 92) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRFSVKIWKIFTDCFNCLPVAALIDERIFCMHGGLSPELKS
7. ( 94) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIRLWKTFTDQFNQLPVAALIDDKILCMHGGLSPELKT
8. ( 94) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDVQRRYNIIRLWKTFTDQFNQLPAQALVDEKILCMHGGLSPELKS
9. ( 94) KQSLETICLLSFKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIRLWRTFTDCFNCLPVAALIDDKILCMHGGLSPELKS
10. ( 93) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYSIKLWKTFTDCFNCLPVAALIDDKILCMHGGLSPELNS
11. ( 96) KQSLEVICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIIDKIFTMHGGLSPDLNS
12. ( 96) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIIDKIFTMHGGLSPELNS
13. ( 97) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIIDKIFTMHGGLSPELNS
14. ( 96) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPVAALIDDKILCMHGGLSPELNS
15. ( 0) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYSIKIWKFTFTDCFNCLPIAAVIDNKIFCCHGGLSPDLHD
16. ( 96) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIVDEKIFCCHGGLSPDLQS
17. ( 96) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIVDEKIFCCHGGLSPDLQS
18. ( 2) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIVDEKIFCCHGGLSPDLQS
19. ( 0) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIVDEKIFCCHGGLSPDLQS
20. ( 0) KQSMETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPVAALVDEKIFCCHGGLSPDLQS
21. ( 2) IQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPVAALVDEKIFCCHGGLSPDLQS
22. ( 95) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPVAALVDEKIFCCHGGLSPDLQS
23. ( 95) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPVAALVDEKIFCCHGGLSPDLQS
24. ( 95) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPVAALVDEKIFCCHGGLSPDLQS
25. ( 97) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIIDKIFTMHGGLSPDLQS
26. ( 0) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIVDEKIFCCHGGLSPDLQS
27. ( 97) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIVDEKIFCCHGGLSPDLQS
28. ( 0) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIVDEKIFCCHGGLSPDLQS
29. ( 97) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIVDEKIFCCHGGLSPDLQS
30. ( 97) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIVDEKIFCCHGGLSPDLQS
31. ( 0) KQSLETICLLLAYKIKYPENFFLLRGNHECASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIVDEKIFCCHGGLSPDLQS
    * * * * *

```

```

1. VEQIKNIKRPTDVPDSGLLCDLLWSDPSKDVKGWGMNDRGVSYTFGPDKVAEFLIKNDMDLICRAHQVVEDGYEFFADRQLVTIFSAPNY
2. VEQIKNIERPTDVPDAGLLCDLLWSDPSKDVKGWGMNDRGVSYTFGADKVAEFLIKNDMDLVCAHQVVEDGYEFFADRQLVTIFSAPNY
3. LDEIRNLPRPTMIPDTGLLCDLLWSDPGKDVKGWGMNDRGVSYTFGPDKVSSEFLTKHDLDLVCAHQVVEDGYEFFADRQLVTIFSAPNY
4. LDEIRNIPRPADIPDHGLLCDLLWSDPKDIEGWGENDRGVSYTFGADKVAEFLQTHDLDLICRAHQVVEDGYEFFANRQLVTIFSAPNY
5. LDDIRRIIRPIDVDPQGLLCDLLWADPDREIQGWGENDRGVSYTFGADKVAEFLQTHDLDLICRAHQVVEDGYEFFAKRQLVTIFSAPNY
6. LRQIRDIRPTDIPDRGLLCDLLWSDPKDVKGWGMNDRGVSYTFGSDIVSGFLKRLDLICRAHQVVEDGYEFFANKQLVTIFSAPNY
7. LEQIRRIIRPTDVPDTGLLCDLLWADPDKDVQGWGENDRGVSYTFGPDVTTEFLQRHDLDLVCAHQVVEDGYEFFAKRQLVTIFSAPNY
8. LDAIRHIPRPTDVPDTGLLCDLLWSDPKDVAGWGENDRGVSYTFGPDVTGFLQKHDLDLICRAHQVVEDGYEFFAKRQLVTIFSAPNY
9. LEQIKRIIRPTDVPDSGLLCDLLWADPDKDIQGWGENDRGVSYTFGPDVTEFLQKHDLDLVCAHQVVEDGYEFFAKRQLVTIFSAPNY
10. LDQIQRIIRPTDIPDTGLLCDLLWSDPEKDLTGWGENDRGVSYTFGADVSRFLQKHDLDLICRAHQVVEDGYEFFGKRQLVTIFSAPNY
11. MDQIRIRPTDVPDTGLLCDLLWSDPKDLTGWGENDRGVSYTFGPDVSRFLQKHDLDLICRAHQVVEDGYEFFGKRQLVTIFSAPNY
12. MEQIRRMVRPTDIPDVGLLCDLLWSDPKDIEGVSENGRGVSFTFGPDVVRFLQKHDMELICRAHQVVEDGYEFFSKRQLVTIFSAPNY
13. MEQIRRMVRPTDIPDCGLLCDLLWSDPKDITGWSENGRGVSFTFGPDVVRFLQKHDMELICRAHQVVEDGYEFFSKRQLVTIFSAPNY
14. MEQIRRMVRPTDVPDTGLLCDLLWSDPKDVQGWGENDRGVSFTFGPDVVRFLQKHDMELICRAHQVVEDGYEFFARRQLVTIFSAPNY
15. FEQIRSLPRPTDVPDTGLLCDLLWSDPEKEIQGWGENDRGVSYTFGNDIVTKFLNRHDLDLICRAHQVVEDGYEFFARRQLVTIFSAPNY
16. MEQIRRMVRPTDVPDTGLLCDLLWSDPKDVQGWGENDRGVSFTFGADVVRFLQKHDMELICRAHQVVEDGYEFFAKRQLVTIFSAPNY
17. MEQIRRMVRPTDVPDTGLLCDLLWSDPKDVQGWGENDRGVSFTFGADVVRFLQKHDMELICRAHQVVEDGYEFFAKRQLVTIFSAPNY
18. MEQIRRMVRPTDVPDTGLLCDLLWSDPKDVQGWGENDRGVSFTFGADVVRFLQKHDMELICRAHQVVEDGYEFFAKRQLVTIFSAPNY
19. MEQIRRMVRPTDVPDTGLLCDLLWSDPKDVQGWGENDRGVSFTFGADVVRFLQKHDMELICRAHQVVEDGYEFFAKRQLVTIFSAPNY
20. TEQIRRMVRPTDVPDQGLLCDLLWADPDKENAGWAENDRGVSFTFGTDINKFLNKHEFDLICRAHQVVEDGYEFFSKRQLVTIFSAPNY
21. MEQIRRMVRPTDVPDQGLLCDLLWADPDKETLGWGENDRGVSFTFGTEVVRFLQKHDMELICRAHQVVEDGYEFFAKRQLVTIFSAPNF
22. MEQIRRMVRPTDVPDQGLLCDLLWSDPKDITMGWGENDRGVSFTFGAEVVRFLQKHDMELICRAHQVVEDGYEFFAKRQLVTIFSAPNY
23. MEQIRRMVRPTDVPDQGLLCDLLWSDPKDITMGWGENDRGVSFTFGAEVVRFLQKHDMELICRAHQVVEDGYEFFAKRQLVTIFSAPNY
24. MEQIRRMVRPTDVPDQGLLCDLLWSDPKDITGWGENDRGVSFTFGAEVVRFLQKHDMELICRAHQVVEDGYEFFAKRQLVTIFSAPNY
25. MEQIRRMVRPTDVPDQGLLCDLLWSDPKDVAGWGENDRGVSFTFGAEVVRFLQKHDMELICRAHQVVEDGYEFFAKRQLVTIFSAPNY
26. MEQIRRMVRPTDVPDQGLLCDLLWSDPKDVQGWGENDRGVSFTFGPEVVRFLQKHDMELICRAHQVVEDGYEFFAKRQLVTIFSAPNY
27. MEQIRRMVRPTDVPDQGLLCDLLWSDPKDVQGWGENDRGVSFTFGAEVVRFLQKHDMELICRAHQVVEDGYEFFAKRQLVTIFSAPNY
28. MEQIRRMVRPTDVPDQGLLCDLLWSDPKDITLGWGENDRGVSFTFGAEVVRFLQKHDMELICRAHQVVEDGYEFFAKRQLVTIFSAPNY
29. MEQIRRMVRPTDVPDQGLLCDLLWSDPKDITLGWGENDRGVSFTFGAEVVRFLQKHDMELICRAHQVVEDGYEFFAKRQLVTIFSAPNY
30. MEQIRRMVRPTDVPDQGLLCDLLWSDPKDVQGWGENDRGVSFTFGAEVVRFLQKHDMELICRAHQVVEDGYEFFAKRQLVTIFSAPNY
31. MEQIRRMVRPTDVPDQGLLCDLLWSDPKDVQGWGENDRGVSFTFGAEVVRFLQKHDMELICRAHQVVEDGYEFFAKRQLVTIFSAPNY
    * * * * *

```


5-12 続き

- 1. CGEFDNAGAMMSVDESLMCSFQILKPADRR (4)
 - 2. CGEFDNAGALMSVDESLMCSFQILKPVDRR (4)
 - 3. CGEFDNAGAMMSVDENLMCSFQILKPAEKK (9)
 - 4. CGEFDNAGAMMSVDDSLTCSFQILKASEKK (23)
 - 5. CGEFDNAGALMSVDDSLTCSFQILKASEKK (23)
 - 6. CGEFDNAGAMMSVSEDLTCSFQILKSNDKK (19)
 - 7. CGEFDNAGALMSVDDTLMQSFQILKPSEKK (20)
 - 8. CGEFDNAGALMSVDDTLMQSFQILKPSEKK (20)
 - 9. CGEFENAGAMMSVDETLMCSFQILKPAETK (5)
 - 10. CGEFDNVGAMMSVNEDLLCSFQILKPAEKR (24)
 - 11. CGEFDNAGAMMSVDESLLCSFQILKPAEKK (26)
 - 12. CVEFDNAGAMMSVDESLLCSFQILKPAQKS (11)
 - 13. CGEFDNAGAMMSVDESLLCSFQILKPAEKK (6)
 - 14. CGEFDNAGGMMTVDDTLMCSFQILKPSEKK (29)
 - 15. CGDYDNAGGMMTVDENLLCSFQILKPSEKK (26)
 - 16. CGEFDNAGGMSVDETLMCSFQILKPSEKK (26)
 - 17. CGEFDNAGGMSVDETLMCSFQILKPSEKK (26)
 - 18. CGEFDNAGGMSVDETLMCSFQILKPSEKK (26)
 - 19. CGEFDNAGGMSVDETLMCSFQILKPSEKK (27)
 - 20. CGEFDNAGAMMSVDESLMCSFQIIKPPEKK (24)
 - 21. CGEYDNAGAMMSVDETLMCSFQILKPQDKK (44)
 - 22. CGEFDNAGAMMSVDDTLMCSFQILKPADKR (27)
 - 23. CGEFDNAGAMMSVDDTLMCSFQILKPADKR (2)
 - 24. CGEFDNAGAMMSVDNTLMCSFQILKPVEKR (2)
 - 25. CGEFDNAGAMMSVDETLMCSFQILKPVDKK (30)
 - 26. CGEFDNAGAMMSVDETLMCSFQILKPADKK (27)
 - 27. CGEFDNAGAMMSVDETLMCSFQILKPADKN (28)
 - 28. CGEFDNAGAMMSVDETLMCSFQILKPAEKK (23)
 - 29. CGEFDNAGAMMSVDETLMCSFQILKPAEKK (21)
 - 30. CGEFDNAGAMMSVDETLMCSFQILKPAEKK (21)
 - 31. CGEFDNAGAMMSVDETLMCSFQILKPAEKK (21)
- * * * * *